Technical Bulletin 162 June 1943

Studies of Semen and Semen Production

R. E. Comstock, W. W. Green,
L. M. Winters, and A. W. Nordskog
Division of Animal and Poultry Husbandry

University of Minnesota
Agricultural Experiment Station

Studies of Semen and Semen Production

R. E. Comstock, W. W. Green,
L. M. Winters, and A. W. Nordskog
Division of Animal and Poultry Husbandry

University of Minnesota
Agricultural Experiment Station

Accepted for publication June 28, 1943

CONTENTS

2M—6-43

Studies of Semen and Semen Production

R. E. Comstock, W. W. Green, L. M. Winters, and A. W. Nordskog

Introduction

THE VOLUME of literature concerning semen and particularly sperm cells has increased greatly within the last 20 years and especially during the last 10. This has resulted partly from a somewhat belated recognition of the fact that the male as well as the female may be responsible for impaired fertility and partly from the recent development of artificial insemination for practical and extensive use.

The Russian scientist Ivanov *(26)* was probably the first to visualize the great contribution to animal improvement that was possible through the extensive use of artificial insemination. This led him to initiate his long series of researches on the physiology of sperm cells and factors relating to the preservation of their viability and fertilizing capacity as well as his work on practical insemination technics. His early discovery *(26)* of the principle of storing semen at reduced temperature was without doubt the greatest single stimulus to the extended use of artificial insemination.

Williams and his co-workers *(59, 60, 70, 71, 72)* were among the first to emphasize the role of the male in infertility. They sought to find means of detecting infertility in otherwise healthy males and were the first to utilize uniformity of shape and size as a criterion of sperm cell quality. Their work has doubtlessly stimulated some of the more recent work on other methods of semen evaluation.

Material and Methods

Sheep were used throughout because lower initial and maintenance costs and greater ease of handling made it possible to work with larger numbers than if another class of farm animals had been chosen. Some of the rams were purebred Shropshires from the experiment station flocks; the rest were grades of mixed Shropshire and Hampshire breeding. Unless specified they were

fed alfalfa hay and sufficient concentrates[1] to maintain reasonable but not high condition. Salt and bone meal were offered ad libitum.

Semen was collected by mating the rams with a diestrual ewe and drawing the semen from the vagina of the ewe with a special glass pipette fitted with a rubber suction bulb. Prior to each collection the vagina was rinsed with the preserving solution for sheep semen described by Winters et al. *(74);* this was done with the collection type pipette. Extreme care was always taken to avoid leaving more than mere traces of the rinsing saline in the vagina. This method of collection[2] had the definite advantage of simplicity. McKenzie and Berliner *(33)* stated, "By this method the whole ejaculate could be collected without leaving significant amounts in the vagina." Terrill *(64)* and Brady and Gildow *(4)* reported no striking differences in quality between semen collected in this way and by the use of an artificial vagina. The latter authors found that less dense semen was obtained by electrical stimulation *(22)*.

Statistical methods, drawn largely from Fisher *(16)* and Snedecor *(63),* were used to assist in evaluation of most of the results obtained. However, a conscious effort was made to include no more statistical detail in reporting work than appeared necessary for clarity. *The word "significant" is used throughout in the statistical sense only* and indicates a probability of 5 per cent or less of occurrence being due to chance alone.

Other materials and technics will be described in the discussion of specific parts of the work.

Semen Evaluation

Efficient criteria of semen quality are essential to numerous laboratory studies, e.g., comparisons of spermatozoa treated differently or produced by males treated differently. They are also needed in the field to test the fertility of males being considered for use in breeding programs or to check periodically the quality of semen produced by males in continuous use. The latter applies particularly to artificial insemination organizations where a large number of females are bred to each sire. Here any delay in recognizing impaired fertility may result in considerable loss.

The quality of semen is dependent upon both a sufficient num-

[1] Concentrate mixture was composed of corn, 80 per cent; oats, 15 per cent; and linseed oil meal, 5 per cent.

[2] Methods more suitable for other classes of livestock or when semen is to be used for artificial insemination are discussed by Green et al. *(21)*.

ber of vigorous sperm cells and a seminal plasma that is favorable to the sperm cells' duration of viability and fertilizing capacity. The quality of the sperm cells themselves depends on their ability to travel up the normal reproductive tract of a female and to fertilize egg cells contacted under favorable conditions. It has never been demonstrated that these two characteristics are necessarily correlated. It is possible that sperm[3] capable of reaching the oviduct and remaining alive until ovulation occurs may not always be capable of normal fertilization, while, conversely, other sperm might have fertilizing capacity but lack viability.

ESTIMATION OF SEMEN DENSITY

The approximate number of sperm cells per unit volume of semen is usually determined with a hemocytometer. This method is sufficiently precise for most purposes but is tedious and slow (five or less samples per hour). A more rapid method of at least equal accuracy is needed.

The error of the hemocytometer was estimated from duplicate counts on 24 different samples. Two counts ranged from 89 to 598 per 0.0001 cu. mm. and averaged 337. The within sample or error variance was 719; the standard error of counting, $\sqrt{719}$ or 26.8. This is in reasonable agreement with the results of Berkson et al. *(3);* they reported 1,324 as the error variance of red cell counts averaging 462.

Measurement of Light Transmission

As the number of cells per unit volume increases, semen becomes more opaque. Burbank *(6)* made use of this when he estimated cell numbers of sperm suspensions by comparing them with $BaSO_4$ opacity standards.

The authors have estimated semen density from light transmission through a standard dilution of semen. Measurements were made with a photoelectric colorimeter of the Evelyn type using a 540 filter.[4] The light source was adjusted to give a galvanometer reading of 100, using sheep semen diluter *(74)* as a blank. The test reading was made after replacing the blank with diluted semen (0.08 ml. in 15 ml. of the diluter). Measurements of light transmission and hemocytometer counts were made on 24 samples of semen.

[3] The terms "sperm," "sperm cells," and "spermatozoa" are used synonymously throughout this publication.

[4] There was no special reason for the choice of this particular filter except that its use resulted in a satisfactory reduction in the amount of light. Another might have done as well.

Analysis of the data revealed a curvilinear relation between galvanometer readings and cell counts. However, the quantity $x = 2$ minus $\log_{10}$ of the galvanometer reading was found to be related to cell count in an essentially linear manner. This is, of course, the transformation used in the determination of dissolved colored substances (in accordance with Behr's law) in standard analytical procedures using the Evelyn colorimeter. It is of some interest to find the same relationship holding for suspended sperm cells. The regression relationship between galvanometer reading and semen density was

$$Y = -19.4 + 1365\ x$$

where Y = an estimate of the number of sperm per .0001 cu. mm. of semen. This relationship is presented graphically in figure 1. It is worth noting that Y equals zero when the galvanometer reading is about 97. This is very close to the reading obtained using seminal plasma containing no sperm cells. The regression of count on x accounted for 95.6 per cent of the variance between samples, the equivalent of a correlation of 0.978 between x and hemocytometer count. The squared standard error of estimate of Y was 745. It is, of course, due in part to errors associated with hemocytometer counts. The difference (26 ± 312) between it and the variance of hemocytometer counts is an estimate of the

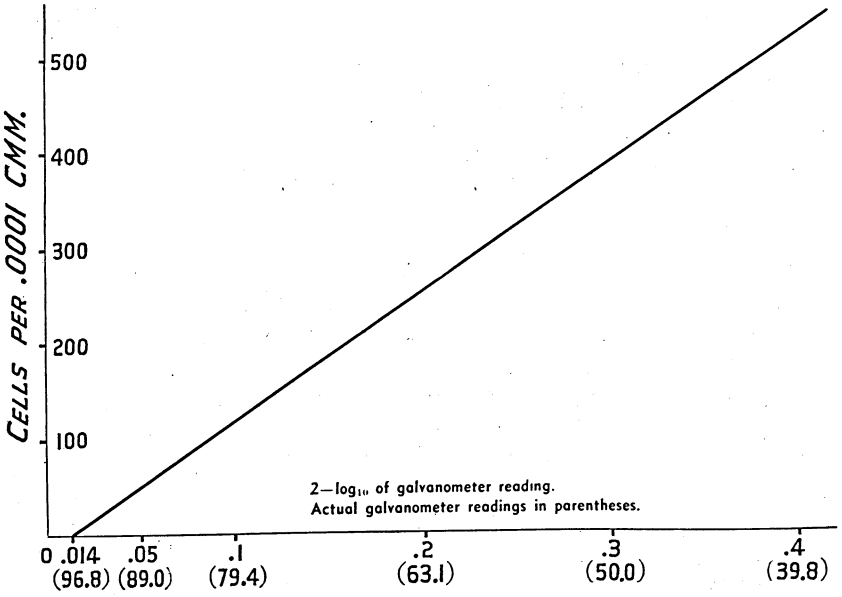

FIG. 1. The relationship between sperm cell count and light transmission as measured with a photoelectric colorimeter

variance of predicting the true cell number from galvanometer readings. Assuming this estimate was as much as two standard errors too low, the standard error of predicting true cell number ($\sqrt{26 + (2 \times 312)} = 25.3$) would still be slightly less than the standard error of the hemocytometer (26.8) reported above. It appears more likely that the colorimeter will estimate semen density as accurately as the mean of two or three hemocytometer counts.

The total time required by this method is no more than that needed to dilute the semen and load the counting chamber when using the hemocytometer. Since the first report of this work (Comstock and Green, *12*), at least two other workers have used the method, apparently with satisfactory results *(24, 27)*.

Comparison with Opacity Standards

The method described above is not practical for most field work since a photoelectric colorimeter would seldom be available. Burbank's method if sufficiently accurate could, however, be very useful in the field. Consequently, an estimate of its accuracy was obtained.

A series of seven opacity standards in which the barium sulphate content ranged from 0.05 to 0.20 per cent at intervals of 0.025 per cent was used. Five ml. of each were put in test tubes 11 mm. in diameter together with a crystal of thymol to prevent bacterial growth. The tubes were stoppered and sealed with paraffin to prevent evaporation. One tenth of a milliliter of semen was mixed with 5 ml. of sheep semen diluter *(74)* in the same type of tube used for the standards. The opacity of this suspension and the barium sulphate suspensions[5] was then compared. This was done in front of an electric light with a piece of window screen immediately back of the tubes. The ease with which the cross wires were seen through the suspensions facilitated comparison.

Twenty-three samples were ranked 1, 2, 3, etc., according to whether suspensions of them corresponded in opacity to the 0.05 per cent suspension of barium sulphate, the 0.05 per cent plus, the 0.075 per cent minus, etc. Two counts were made on each sample using the hemocytometer; these were 23 of the 24 samples on which error of the hemocytometer was determined. The variance of estimating the mean of two counts from opacity rank was 1,164. Subtracting 719/2 (the error variance of the mean of

[5] The suspensions tend to settle out while standing; hence it is always necessary to shake them vigorously before comparisons are made.

two counts) leaves 805. An estimate of the standard error of predicting true cell number from opacity rank is therefore $\sqrt{805}$ or 28.4. Thus it appears that opacity rank may be approximately as accurate as the hemocytometer for the determination of semen density.

Extreme care must be exercised in the preparation of barium sulphate suspensions if results from different series of standards are to be comparable. The most satisfactory method of preparation is the precipitation of barium sulphate by mixing solutions of barium chloride and sodium sulphate. For most field work a series of no more than five standards is sufficient; it should include 0.06, 0.09, 0.12, 0.15, and 0.18 per cent suspensions. Detailed directions for preparation of such a series are as follows:

1. Prepare (a) an aqueous solution containing 321.2 mg. of barium chloride in 100 ml. and (b) an aqueous solution containing 219.0 mg. of sodium sulphate in 100 ml. (These solutions must be prepared in 100 ml. volumetric flasks using distilled water.)
2. Into a series of five tubes in which standards are to be prepared, pipette 1.0, 1.5, 2.0, 2.5, and 3.0 ml., respectively, of the barium chloride solution.
3. Add 1.0, 1.5, 2.0, 2.5, and 3.0 ml., respectively, of sodium sulphate solution to the tubes containing the corresponding amounts of barium chloride solution. The sodium sulphate solution must be allowed to run in rapidly from a pipette calibrated "to deliver" since the rate of addition of this solution has a measurable effect on the opacity of the suspensions.
4. Allow the suspensions to stand 30 minutes. Then add sufficient distilled water to the first four to bring them to 6 ml. total volume. The fifth will already contain 6 ml.

The original solutions of barium chloride and sodium sulphate must be prepared at least four days before the suspensions are prepared and the suspensions should be made up three or four days before they are first used.

A shortcoming of the above method is that each laboratory or worker using it to estimate absolute values in contrast to using it as a relative measure will need to standardize a prediction curve since comparisons with the standards will not be made in exactly the same way by any two workers. However, for the approximation of relative values the method could be of considerable use in the field. Very little time is required and the equipment is inexpensive.

Anyone with a certain amount of experience can estimate density fairly accurately without standards for comparison, but the average level of their estimates will shift from time to time in the absence of a standard for reference. For example, there is marked seasonal variation in the average density of ram semen (the same has been reported for bull semen [*15*]). As a result, samples of equal density may appear relatively much more dense in one season than in another. The barium sulphate standards serve to anchor one's judgment to about the same range at all times.

CRITERIA OF SPERM CELL QUALITY

Literature

Motility is a necessary attribute of sperm; cells not motile under favorable conditions are known to be physiologically nonfunctional. It is natural, therefore, that degree of motility should suggest itself as a measure of quality. However, while motility estimates have been widely used and there is evidence that they have some value *(14, 22, 23, 50, 68)*, they have been found inadequate by many workers *(33, 40, 46, 70)*. This has led some to state that motile spermatozoa are not necessarily fertile. This may be true but is misleading. Though insemination, either natural or artificial, with motile sperm may fail to produce conception, it is probable that in many such cases it is motility that fails and not some other attribute of the sperm as the above statement implies. Asdell and Salisbury *(1)* moved the testis and epididymis of male rabbits from the scrotum to the body cavity by operative procedure. Spermatogenesis ceased within 25 hours. While motile sperm were ejaculated up to the fourteenth day following operation, fertilizations were not obtained in matings made after the eighth day. They found that in those matings sperm did not reach the oviduct of the female. Though motility was present at the time of mating it appears very likely that loss of motility was responsible for failure to conceive.

The minimum motility duration required of spermatozoa is that needed to travel from the vagina to the upper end of the oviduct. This has been variously reported for sheep. Schott and Phillips *(61)* and Phillips and Andrews *(53)* are in substantial agreement; the former report about 20 minutes, the latter a minimum of 30 minutes. Another group in equally as good agreement *(20, 28, 55)* all report 5 or 6 hours are required. Brewster et al. *(5)* concluded that in heifers 4 to 7 hours and in mature cows 6 to

9 hours are needed for sperm to reach the infundibulum. Sperm with the minimum viability necessary would doubtlessly be of lower quality than those used by the above workers and would, therefore, require still more time to ascend the female tract.

Time of ovulation also affects the requirement for motility duration. This is particularly true in the bovine when ovulation may occur anywhere from 0-20 or 30 hours after the cessation of estrus *(5, 51)*, while mating or insemination usually occurs during estrus.

Obviously sperm must not only be motile at the time of insemination but must also have the power to remain motile for some time thereafter. Estimates of the motility of fresh sperm are not closely related to duration of motility. McKenzie et al. *(34)* reported that samples of boar semen all highly motile when fresh varied from one to 24 days in duration of motility when stored. A correlation of only 0.3 between estimated initial motility and motility after 6 days of storage was found at Minnesota. The fresh sperm of one particular ram was invariably highly motile yet never maintained more than a very low order of motility for even as long as 24 hours.

The inadequate prediction of duration from estimates of initial motility may result from (a) inability to distinguish intensities above a certain level, or (b) low correlation between true initial intensity and duration of motility. It is possible, though there is no specific evidence, that initial motility is a better criterion of viability in the female reproductive tract than in storage. On the contrary, Herman and Swanson *(25)* reported that bull semen which effected conception averaged longer life in storage than semen that did not; the difference in initial motility was negligible and actually in the opposite direction.

Lasley et al. *(32)* reported a staining method for differentiating live and dead spermatozoa in ram semen. They also found *(31)* the proportions of sperm in bull semen stained by this method after temperature shock (10 minutes at 0° C.) and after 144 hours in storage were directly related. They did not state the magnitude of the correlation nor did they show the stain to be as specific for live sperm in bull semen as in ram semen. MacLeod *(39)* reported that the method does not distinguish live from dead cells in the semen of humans.

Numerous workers have presented evidence that the proportion of morphologically abnormal sperm cells in semen is a useful criterion of male fertility. Williams *(70)* and Williams and Savage *(71, 72)* originated the method in work with bulls. The first

applications to human fertility were by Moench *(42, 43)* and Mason *(40)*. Moench *(44, 45, 46, 47)* has published persistently on the subject and has done much to increase the emphasis placed on semen evaluation in the diagnosis of human sterility. The method has also been applied to rams, boars, and stallions (McKenzie and Phillips, *35, 36,* Phillips, *52,* and Williams, *69*). Other literature is cited by McKenzie and Berliner *(33)*.

Williams and Savage and Lagerlof *(29)* are in substantial agreement on the maximum number of abnormally formed sperm cells found in the semen of bulls with good breeding records. The former workers report no more than 17 per cent while the latter found usually no more than 18 but occasionally as many as 20 per cent. On the other hand, Phillips found no more than 17 per cent abnormal sperm in the semen of fertile boars, whereas Milowanow (cited by McKenzie and Berliner) reported as many as 30 per cent. Such differences can reasonably be expected. Doubtlessly some technicians classify cells as abnormal that others would call normal. This tendency could be reduced to a minimum if all would follow the same standard, e.g., by use of the types illustrated by Lagerlof *(29)* or Moench and Holt *(47)*. Differences among workers may also arise from variation in the technic of preparing the material; this was clearly demonstrated by Salisbury, Willett, and Seligman *(58)*. In addition, contrary to the impression given by some of the literature, there is probably no distinct line of demarcation between animals of good and poor fertility. Herman and Swanson *(25)* recorded data on 55 bulls with known breeding records. In general their results verify the value of sperm cell morphology for evaluating fertility. However, there was one bull of four with "questionable" breeding records that had an average, in 32 samples, of only 9 per cent abnormal sperm and one of five with "poor" records that averaged only 13 per cent in 25 samples observed. Among 42 with "good" records there were 9 that averaged from 20 to 33 per cent abnormals. Forty-one samples were examined from one of these but no more than three from any of the others.

There is no general agreement on the most serious types of malformations. Moench and Holt *(47)* stress the importance of head abnormalities and particularly of narrow or tapering heads. They found that tail and middle piece aberrations had no bearing on fertility. This point has not been emphasized by other workers. The data of Herman and Swanson *(25)* fail to confirm it in the case of bull semen. McKenzie and Phillips *(36)* laid particular emphasis on tail and middle piece abnormalities. There

is considerable disagreement regarding the so-called protoplasmic drop. One group of workers believes their presence on sperm cells indicates immaturity, while another finds no evidence to support that belief (see McKenzie and Berliner, *33*).

The value of sperm morphology apparently does not result from correlation with viability. If it did, the proportion of abnormal cells should increase with age of sample. Williams and Savage *(72)* state, "In other cases we have allowed samples to remain in the laboratory for several days after the original examination had been made. Decomposition became marked, yet smears from such material showed approximately the same proportion of degenerate heads as the fresh sample." Correlations between the per cent of malformed sperm and motility duration in storage were calculated from the data of Herman and Swanson *(25)*. A correlation (based on 206 samples from eight animals) of 0.19 was found within bulls. If viability increases as aberrant forms decrease, the figure should at least be negative. The averages of 45 bulls, weighted according to the number of samples from each bull, yielded a correlation of —0.25—negative but very low.

Another method of evaluating sperm cell quality is based on biometric constants calculated from measurements of sperm head lengths. It is probably closely allied to the enumeration of abnormal forms since it is reasonable that variation in the shape of sperm heads should be reflected in a series of measurements of their lengths. Williams and Savage *(71, 72)* and Savage, Williams, and Fowler *(59)*, working with bulls, found that both coefficients of variation over 3.5 or 4.0 and skew distributions of head lengths were strongly indicative of infertility. The same authors *(60)* working with stallions reported similar findings, though the information on actual breeding performance was not as complete as desired.

Moench *(44, 45)*, Moench and Holt *(48)*, and Mason *(40)* studied head length of human sperm. Again high coefficients of variation (over 11.5 to 12.0) and skew distributions were found to indicate low fertility.

Comstock and Brady *(11)* working with ram semen found high coefficients of variability to be associated with low duration of motility when stored. They also found indications that when the mean head length varied too much from an optimum that motility duration was reduced. However, their data were not extensive enough to furnish accurate estimates of the magnitudes of these relationships.

Work at this laboratory on the evaluation of sperm cell quality has, for the most part, concerned (a) the relationship of sperm viability to rates of glycolysis and respiration and (b) sperm cytology with special emphasis on a specific structural detail of the sperm head.

Glycolysis and Respiration of Sperm Cells

Glycolysis as used herein refers to the production of acid from glucose or other substrates by the action of cell enzymes. The acid formed is frequently considered to be lactic but the authors are using a broader definition because the method employed to measure acid production is nonspecific. Respiration refers to those reactions which result in oxygen uptake by the tissue. The literature concerning the relationship of these reactions to motility of sperm was reviewed by Comstock *(9)* and more recent publications are cited by Moore and Mayer *(49)*. It has been demonstrated that neither glycolysis nor respiration is indispensable to motility of the sperm cells of some species provided the other is allowed to proceed. However, MacLeod *(37, 38)* reported that motility depends on glycolysis in the case of the sperm of humans. Unimpaired motility has not been demonstrated in the absence of both reactions in any species. Lardy and Phillips *(30)* reported observing a low order of motility of bull sperm in circumstances where glycolysis and respiration were either absent or proceeding only at a very slow rate. The relationship of these reactions to physiology of sperm motility suggested that measurements of the rate of one or both might prove useful in evaluating viability. Some work has already been reported from this laboratory *(8, 9, 12)*.

The method of Warburg *(66)* was used to measure both glycolysis and respiration. All glycolysis measurements to be discussed were made anaerobically (under nitrogen) in the presence of glucose according to procedure given by Comstock *(9)*. Respiration was measured under air; sperm cells, with one exception, were suspended in same solution used in measuring glycolysis.

Preliminary tests demonstrated that glycolysis rates were correlated rather strongly with estimated motility but that differences in glycolysis between rams were not entirely accounted for by estimated motility and cell number *(9)*. It has also been shown that both respiration and glycolysis rates decline with the age of stored sperm *(8, 9)*. Both of these facts lend credence to the possibility of using one or the other of these reaction rates as a measure of viability.

EXPERIMENTAL—Glycolysis was investigated before respiration because: (a) Redenz *(57)* had shown that anaerobic motility is supported by glycolysis, and (b) preliminary work on a series of eight samples had shown that during storage, respiration declined much more rapidly than glycolysis, while motility remained strong *(8)*.

Glycolysis rates of 12 samples from each of six rams were measured using 0.1 ml. of semen for each determination. Readings were made at 40- and 80-minute intervals after the initial reading. Motility duration (in hours) was determined on a portion of each sample as described by Comstock *(9)*. Hemocytometer counts of cell numbers were made on all samples. Because the Warburg apparatus used had six manometers it was convenient to handle groups of six samples, one from each ram, in parallel. Variance and covariance of the three quantities measured were analyzed as indicated by the following form:

Source of variation	Degrees of freedom
Rams	5
Replications	11
Rams x replications	55
Total	71

Variance and covariance within rams were separated into parts associated with "replications" and "rams x replications," respectively, because no provision had been made for blank readings to compensate for the effect of variations in the rise or fall of room temperature during the course of the determinations. Thus correlations involving glycolysis which included variation between replications might not have been representative of the actual relationships.

There were revealed significant differences between rams in each of the quantities measured. Correlation coefficients are presented in table 1. The deviations of motility duration from regression on glycolysis rates (in terms of the standard error of estimate) are compared in table 2 with the total variation in motility duration (in terms of the standard deviation). Prediction of motility duration from glycolysis rate and sperm number was not significantly more accurate than prediction from glycolysis alone. Table 3 contains an analysis of deviations from regression, indicating that significant differences in motility duration independent of variation in glycolysis rates existed between the rams. It should be noted, however, that the sum of

Table 1. Correlation Coefficients between Glycolysis (x), Motility Duration (y), and Semen Density (z)

Source of variation	Degrees of freedom	Correlation coefficients		
		r_{xy}	r_{zy}	r_{xz}
Rams	4	.997	.666	.628
Rams x replications	54	.643	.315	.616
Rams plus rams x replications	58	.920	.596	.625

Table 2. Variation in Motility Duration

Source of variation	Standard deviation	Standard error of estimate from regression on glycolysis rate
Rams x replications	61.0	47.1
Rams plus rams x replications	142.1	56.4

Table 3. Analysis of Deviations of Motility Duration from Regression on Glycolysis Rate

Source of variation	Degrees of freedom	Sum of squares	Squared standard error of estimate	F
Rams	5	67,605	13,521	6.1*
Rams x replications	54	119,828	2,219	
Rams plus rams x replications	59	187,433		

* Indicates a probability of less than .01 that an F value as large as this would occur as a result of sampling errors alone.

squares of motility duration resulting from differences between rams was reduced 93 per cent by the elimination of the portion associated with variations in glycolysis.

The above analyses were made using the 80-minute glycolysis determinations. Similar analyses using 40-minute measurements gave results in such close agreement that 40-minute intervals were used in all later work.

Walton and Edwards *(65)* reported respiration rates of sperm cells to be related to the fertility of bulls producing them. At the same time they recognized reasons for believing that glycolysis rates might be more closely related to fertilizing capacity. An attempt was made to determine whether motility duration was more accurately predicted by one than by the other. Rates of glycolysis and respiration, sperm number, and motility duration were obtained for each of 72 samples, six from each of 12 rams. Sperm number per unit volume was estimated in this experiment by the photoelectric colorimeter as described earlier. The rams were divided into four groups of three, and samples from the three rams of a group were run in parallel. Before conducting

this experiment an additional unit had been added to the Warburg apparatus to record a blank reading resulting from any failure of temperature control. The form taken by the statistical analysis was as follows:

Source of variation	Degrees of freedom
Rams	11
Replications within groups	20
Within replications within groups	40
Total	71

The sums of squares and products associated with replications were separated from the remainder within rams in order to get the best possible estimate of the correlation between respiration and glycolysis rates. Any error in the blank determination would have opposite effects on the measurement of glycolysis and respiration rates of all three samples of a replication, raising one and lowering the other. The result would have been to lower the apparent correlation if the above precaution had not been taken.

Correlation coefficients are presented in table 4. The correlations between glycolysis and respiration indicate that the precaution taken regarding replications was justified. While further analysis showed significant differences between rams in both glycolysis and respiration rates which were independent of variation in the other, the high correlation (0.88) between the two for "total minus replications within groups" leads one to doubt that one of them would predict either motility duration or actual fertility much better than the other. Actually the correlation of motility with respiration rate (r_{wy}) was slightly higher than with glycolysis (r_{xy}) in this experiment. However, Hotelling's test *(2)* showed the difference was significant only in the case of the coefficients for "total minus replications within groups." Multiple regression of motility duration on both glycolysis and respiration was significantly better than regression on glycolysis alone but not better than on respiration alone.

Table 4. Correlation Coefficients between Glycolysis (x), Respiration (w), and Motility Duration (y)

Source of variation	Degrees of freedom	Correlation coefficients		
		r_{xw}	r_{xy}	r_{wy}
Rams	10	.937	.756	.859
Replications within groups	19 } 59	.569 }	.492	.427
Within replications within groups	39 }	.735 }		
Total	70	.803	.640	.713
Total minus replications within groups	50	.881	.659	.768

As in the previous experiment, prediction of motility duration was not improved by use of sperm number.

Table 5 contains a comparison similar to that of table 2 between total variation of motility duration and the deviation of motility duration from regressions on the reaction rates. The standard errors of estimates were much the same as in the previous experiment. The lower correlations between motility duration and reaction rates appear to be the consequence of lower total variation. The variations associated with replications were not eliminated in obtaining the errors of estimate as was done in the first experiment. In that case, errors which could have been eliminated by proper procedure would have been introduced by inclusion of replication variation, whereas in this case the equipment had been improved and whatever error remained in blank determinations would always confront one in the use of these reaction rates.

As in the previous experiment, a large portion of the sum of squares for differences in duration of motility between rams was associated with variation in reaction rates, although a significant portion was not.

It appeared possible that the high correlation obtained between glycolysis and respiration rates might have been the result of the amount of substrate available for respiration being dependent on the rate of glycolysis. In another experiment an amount of sodium lactate equal in lactate ion content to one per cent of lactic acid was added to the suspension fluid for use in the measurement of respiration. This eliminated the possibility of substrate limitation affecting the respiration rate; respiration proceeds unhampered when sodium lactate is the only substrate furnished.

Rates of glycolysis and respiration of the fresh sperm, glycolysis rate after four days of storage, sperm number, and estimated motility after four days of storage were recorded for each of 60 samples of semen, four from each of 15 rams. The sperm used for measurements of glycolysis after storage were washed by suspending them in a quantity of the same fluid used for

Table 5. Variation in Motility Duration

Source of variation	Standard deviation	Standard error of estimate from regression on	
		Glycolysis	Respiration
Within rams	49.4	43.4	45.0
Total	78.4	60.6	55.4

dilution in the manometers, centrifuging, and pipetting off the supernatant fluid. Two repetitions of this process have been found sufficient to guard against bacteria accumulated during storage having a perceptible effect on measurement of glycolysis rates. Sperm numbers were estimated by the colorimeter method. Motility at four days was substituted for motility duration to save time. Previous work *(8)* indicated that correlations of glycolysis and respiration with motility at four days would be lower than had been obtained with motility duration. However, there appeared little reason to believe they would be less useful for comparing the value of the two reaction rates in predicting livability. Glycolysis rates were measured at four days to provide another measure of viability.

In this experiment a specific form of replication was not used. When gathering sets of three samples to test at one time, the choice of rams was random except that four samples were tested from each of the 15 used, and no more than one sample from a particular ram was ever included in a set. The form of statistical analysis was as follows:

Source of variation	Degrees of freedom
Rams	14
Within rams	45
Total	59

Correlation coefficients are presented in table 6. The correlation between glycolysis and respiration rates was nearly the same as before, though in this case it was not convenient to eliminate replication variation which would very likely have raised it. Again significant differences in each of the reaction rates independent of variations in the other existed between rams.

The correlations between four-day motility and glycolysis and respiration, respectively, were little different. In this case the correlation with glycolysis was the larger though not significantly so. Prediction of motility at four days on the basis of both reaction rates was significantly better than by respiration alone

Table 6. Correlation Coefficients between Glycolysis (x), Respiration (w), Motility after Four Days' Storage (y), and Glycolysis after Four Days' Storage (x^1)

Source of variation	Degrees of freedom	Correlation coefficients				
		r_{xw}	r_{xy}	r_{wy}	r_{xx^1}	r_{wx^1}
Rams	13	.927	.473	.409	.663	.620
Within rams	44	.707	.546	.468	.432	.441
Total	58	.871	.494	.422	.608	.582

but not better than by glycolysis alone. This is exactly contrary to what was found in the preceding experiment. In view of the high correlation between respiration and glycolysis a logical explanation for this seeming contradiction in results is that measuring both reactions probably accomplishes about the same as would duplicate measures of either one. If the duplicate measurements were considered as two separate sets, it would not be surprising if prediction on the basis of both would be significantly better than on the basis of the set which happened in that particular case to be the least accurate but, though better than on the basis of the other set, not significantly so.

As before there were differences between rams in motility at four days which were independent of variation in respiration or glycolysis.

There was very little difference between the correlations of glycolysis at four days with initial rates of glycolysis and respiration, respectively. Significant differences in four day glycolysis remained between rams after elimination of the differences associated with initial respiration or glycolysis rates.

DISCUSSION—In every case the semen of different rams varied significantly in reaction rates. Earlier work *(9)* had shown such variation to be in part independent of cell numbers and estimated motility. There is a moderately high correlation between motility duration and rates of glycolysis and respiration. As sperm cells age in storage, reaction rates decline. These facts together with the results of Walton and Edwards *(65)* are evidence for the value of reaction rates as viability criteria for ram and bull sperm. It is noteworthy that Chang and Walton *(7)*, using respiration rates to evaluate sperm cells subjected to different temperature sequences, secured very orderly results.

There is no evidence that either glycolysis or respiration is superior to the other for evaluating the viability of ram sperm. This is apparently not true for sperm of all species; MacLeod *(37, 38)* reported that in human sperm, motility is closely related to glycolysis but not to respiration which he states is always insignificant in amount. However, the results of Walton and Edwards *(65)* and the fact that bull sperm, like ram sperm, exhibit relatively high respiration suggest that the value of respiratory rates is comparable in these species.

There were differences between rams in both motility duration and glycolysis rate after four days' storage which were independent of initial reaction rates. The conclusion is clear that factors other than initial reaction rates affect persistence of

motility in storage and that livability can be measured directly more accurately than it can be predicted from glycolysis or respiration. An advantage of reaction rates is the fact that they give estimates when the semen is still fresh, while actual measurement of motility duration is completed only when the sperm are at least partially exhausted. It is also reasonable to question whether motility duration is a better criterion of viability in the female reproductive tract than are reaction rates; different qualities may be necessary for survival in the female than for survival in storage at reduced temperature. It is apparent that the exact characters to be evaluated affect the relative efficacy of the methods available.

Winchester and McKenzie *(73)* conclude that equal cell concentration must be used when comparing absolute respiration rates since, other things being equal, respiration per sperm cell is inversely related to density of the suspension used in measuring the reaction. The situation is similar with respect to glycolysis *(8, 56)*. However, both reaction rates (considered as estimates of viability) and number of sperm affect the quality of semen. Rates per unit volume of semen were used in the work reported above because they are weighted by both the metabolic activity and the number of sperm cells. At the same time they do not require dilution to a standard concentration which is time-consuming and certain to be a source of error.

Cell numbers were positively correlated not only with reaction rates as measured but also with motility duration (table 7). It was not surprising, therefore, that consideration of sperm number did not significantly improve the prediction of motility duration from reaction rates. This implies that prediction from rates per unit number of cells at a standard density would not be as accurate as from rates per unit volume. However, there is a possibility that the measurement of motility duration was biased in favor of the more dense semen. This could result from a tendency for denser samples to appear more highly motile or from an

Table 7. Correlation Coefficients of Sperm Number per Unit Volume (z) with Glycolysis (x) and Motility Duration (y)*

Experiment	Simple correlations	
	r_{xz}	r_{zy}
1	.536	.567
2	.666	.493
3	.629	.337

* Motility duration based on four days of storage in the case of experiment 3.

advantage of density during storage (all stored semen was diluted an equal amount). The latter appears least likely since ram semen was found to retain motility equally as well at a one to three dilution as at the one to one dilution used in this work (unpublished data). On the other hand, there may be a real correlation between cell number and viability of the individual cells. It is not illogical to postulate that as the physiological state of testes improves they would tend to produce both more and higher quality sperm cells. This possibility receives support from the existence of a small positive correlation between sperm number and glycolysis per unit volume divided by sperm number, whereas the general effect of increasing the density of uniform quality sperm is to reduce glycolysis per unit number. The values of this correlation in the three experiments reviewed and in the data of an additional group of 92 samples were 0.04, 0.11, 0.11, and 0.07, respectively.

Walton and Edwards *(65)* report a somewhat higher relation between fertility and respiration rate per unit volume of semen divided by sperm number than between fertility and rate per unit volume. The data were, however, hardly sufficient to establish this beyond reasonable doubt. Furthermore, rate per unit volume should give some measure of reproductive capacity as well as of the quality of the individual sperm produced since it is correlated with sperm number. There is no indication in their report of reproductive load being sufficient to affect the fertility of the bulls with which they worked.

Studies of Sperm Cell Cytology

One of the authors sought to find a variable cytological feature associated with the physiologic state of sperm cells *(17)*. Using aceto-carmine stain in a manner described in his original paper and examining the cell microscopically with the aid of dark field illumination, he observed a small hyaline vesicular structure as a part of the cell membrane at the anterior portion of some sperm (see figure 2). He termed this feature a "vesicle." Considerable evidence has been compiled to show that the pro-

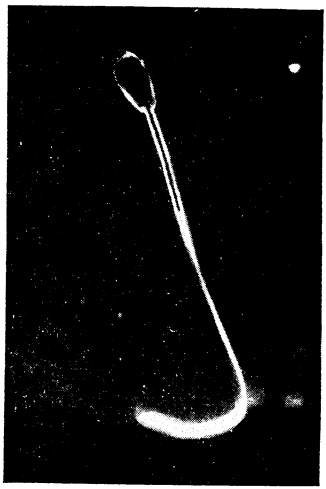

FIG. 2. A ram spermatozoon with a typical "vesicle"

Table 8. Per Cent of Sperm with Vesicles at Different Ages of Stored Samples

Age of samples in days	Ram from which sample was secured			
	54	36	67	32
0	67	51	39	35
1	43	31	24	22
2	42	28	18	21
3	17	7	5	11
4	4	3	5	4
5	0	1	1	1

portion of cells exhibiting a vesicle is related to their average quality. Although all work to be reported was with ram sperm, a comparable structure was found on bull sperm and a somewhat modified one on the sperm cells of boars.

As spermatozoa age in storage, the per cent of cells with vesicles decreases. Typical data are presented in table 8. The relationship between age and number of vesicles is striking. Each figure is based on the observation of 150 cells which were morphologically normal in all other respects.

Vesiculated cells were also counted in the 60 samples used in the last experiment of the previous section, both when fresh and after four days of storage. The condensed data are presented in table 9. Statistical analysis revealed highly significant differences between rams and between fresh and aged samples as would be expected from casual observation of the data. It was interesting, on the other hand, that the interaction of age of sample with individuals from which they were secured was not significant. This implies that regardless of initial level the rate of decline was not uniformly different for one ram than for another. It was a little surprising not to find significant differences in this respect among as many as 15 rams. The same interaction was highly significant in the case of glycolysis rates.

Table 9. Per Cent of Sperm with Vesicles (Average of Four Samples) in Fresh Samples and in the Same Samples after Four Days of Storage

Ram	Age of sample, days		Ram	Age of sample, days	
	0	4		0	4
107	23.5	9.8	77	32.8	14.5
109	24.3	12.3	33	39.0	15.3
112	14.0	9.0	54	29.8	7.8
113	5.8	3.8	N	22.8	10.8
114	9.0	4.0	61	15.5	5.8
115	11.0	4.5	67	24.0	13.5
108	31.5	12.3	101	13.0	4.5
110	20.3	7.0			

Table 10. Per Cent Cells with Vesicles before and after Samples Were Held at 37° C. for 50 Minutes

Sample	Before	After
1	18	0
2	19	1
3	15	2
4	15	2
5	13	1
6	22	5

Under conditions where loss of viability is much more rapid than in storage at reduced temperatures, the decline in per cent cells with vesicles is also more rapid. Thus when vesicles were counted at the beginning and end of 50 minutes at a temperature of 37° C. the decline was comparable to that during a number of days at 4° C. Table 10 contains typical results. Each of these samples was collected from a different ram.

An attempt was made to find whether vesiculated cells ascend the female reproductive tract in greater numbers than other cells. Two ewes were inseminated with samples in which vesicles had been counted, and were killed 12 hours later. The sperm were immediately washed from the infundibulum and vesicular counts made on those sperm which had traveled the entire length of the reproductive tract (table 11). In each case the proportion of vesicles was significantly higher in sperm cells flushed from the infundibulum than in the original sample. Although data from the use of more animals would be desirable, the results strongly suggest that the vesiculated cells are better adapted than others to travel in the female tract.

It was found in three trials that when ejaculates were collected at approximately hourly intervals from the same ram that the proportion of vesiculated cells in the successive samples gradually decreased. There is thus a possibility that lack of the vesicle indicates immaturity of sperm as well as loss of function through age.

While a vesicle was never observed on a cell exhibiting any

Table 11. Comparison of Per Cent* Cells with Vesicles in Samples Used for Insemination and among Sperm Which Reached the Infundibulum in 12 Hours

Ewe	Per cent cells with vesicles	
	In original sample	At the infundibulum
1	9.0	30.0
2	6.0	33.0

* 150 cells observed to obtain each figure.

type of abnormal head, the proportion among sperm cells with tail abnormalities was not different from among sperm with no abnormalities of any sort. Among 23,500 cells observed in one series, 19.66 per cent of those with only tail abnormalities exhibited vesicles, while the per cent among those with no visible abnormalities was 19.94. In another series of 12,500, the corresponding proportions were 16.51 and 16.23, respectively.

Vesiculated cells were counted on all of the fresh samples of the second and third experiments described in the previous section. The relationship between per cent vesicles in fresh samples and livability in storage was low. In the second experiment the partial correlation between per cent vesicles and motility duration with cell number held constant was statistically significant but so low that vesicle number would have little predictive value. Only in the case of the within ram variation did the data on vesicles and glycolysis predict motility duration significantly better than glycolysis alone. Even here the amount by which the standard error of estimate was decreased was so small as to be of little practical importance. While all of the data have indicated a positive relationship between numbers of vesicles and glycolysis rates, this relationship is not high. Correlation coefficients involving glycolysis, vesicle numbers, and sperm number are presented in table 12. Correlations with sperm number are given because vesicles were counted per unit number of cells, while glycolysis was measured on unit volumes of semen.

The positive relationships of vesicle numbers to glycolysis rates and motility duration are so low as to suggest that they result only from the fact that all three decrease on the average with age of cells rather than from a direct physiological relationship.

While the evidence does not indicate a high relationship between vesicle number and viability, the data of table 11 suggest that this structure may have an important function relative to ascent of the female reproductive tract. It is probable that qualities other than viability are necessary to ascent of the female tract by sperm; such a factor suggested by Miller and Kurzrok

Table 12. Correlation Coefficients Involving Glycolysis Rate (x), Per Cent Vesicles (u), and Sperm Numbers (z)

Experiment	Partial correlations	Multiple correlations	Simple correlations	
	$r_{xu.z}$	$r_{x(uz)}$	r_{xu}	r_{xz}
2	.375	.722	.294	.666
3	.170	.642	.382	.629

(41) is the ability of sperm cells to dissolve normal cervical mucus. It is entirely possible that the "vesicle" is an indicator of, or is itself one of, those qualities. Thus, regardless of its relationship to viability it should on the basis of present evidence be considered as a possible important characteristic of high quality sperm.

The report of Seymour and Benmosche *(62)* is interesting in relation to the vesicular structure. Photographs of human sperm taken through an electron microscope revealed an irregularity at the vertex of the sperm head which they suggest may have a special function in fertilization of the egg. It is not impossible that what appeared as a "notch" in their photograph is the equivalent in the human of the "vesicle" discussed above in the ram and bull.

Relationship of Semen Characteristics to Probability of Conception

An attempt was made to compare the efficacy of various measures for predicting the ability of semen to induce conception when used for insemination. Sixty-five ewes were employed. In order to obtain a complete record of their estrus cycles they were "teased" three times a day, using a ram prevented from actual copulation by an apron tied securely around his middle. No ewe was inseminated during the first estrus period observed; knowledge of the length of the first estrus made it possible within limits to inseminate all ewes at a standard time relative to the end of the heat period. Each ewe was inseminated with 0.3 ml. of freshly collected semen. The remainder of each sample used was taken immediately to the laboratory where slides were prepared for cytological studies, glycolysis rates measured, and sperm numbers estimated by means of the photoelectric colorimeter. Ewes which did not conceive were inseminated in as many as four successive heat periods. All which had not then conceived were mated by direct service with a highly fertile ram. Subsequently all but three ewes were slaughtered. Number of corpora lutea and number of embryos were recorded if fertilization had occurred. Two of the five ewes not pregnant exhibited gross abnormalities of the reproductive tract. The records of all inseminations involving the five ewes which did not finally conceive after natural mating were eliminated from consideration.

Table 13 contains the means of various criteria for samples that induced fertilization and those that did not. None of the differences were significant although the samples in group 1 were on the average slightly superior in each item considered.

Table 13. Comparison of Samples Inducing Conception with Those Not Inducing Conception

Criteria	Group I: Samples inducing conception N = 45	Group II: Samples not inducing conception N = 47	Diff. S.E. of Diff.
Glycolysis per 0.1 ml. semen	32.3	29.0	0.63
Glycolysis per unit number of sperms*	0.1343	0.121	1.05
Sperms per 0.0001 cu. mm. semen	247.9	230.1	0.84
Vesiculated sperm in 500	14.7	14.4	0.17
Sperm with normal form in 500	374	361	0.63

* Glycolysis per unit volume divided by number of sperm per 0.0001 cu. mm. semen.

On the average, there were more factors other than semen quality operating to reduce the probability of conception in the ewes that did not conceive in the first estrus period than in those that did. Consequently, whether or not conception resulted in subsequent inseminations seemed to depend more on factors other than semen quality than was the case in the first round. In addition, the semen used for inseminations subsequent to the first averaged higher in all items measured. Because there is some indication that the probability of conception was not greatly different with the exception of the very poorest samples used (see table 14), the result of increased quality of semen may have been practically to equalize the probability of conception resulting from most of the samples used in other than first inseminations. In that case the occurrence of conception would depend more than ever on factors other than semen quality. These considerations appear to justify treating data from initial inseminations separately. Table 15 contains results for this group of data comparable in form to table 13. While the differences in general are larger, still none are significant.

The nature of these data is a complicating factor in their analysis. Whenever factors other than semen quality (of which there conceivably may be many) are responsible for failure of an

Table 14. Proportions of Conceptions Resulting from Semen with Different Levels of Glycolysis

Glycolysis rate (per 0.1 ml. semen)	Number of conceptions	Number failing to conceive
0- 9.9	5	10
10-14.9	5	5
15-19.9	8	6
20-29.9	8	9
30-44.9	7	6
45-59.9	6	5
60 and over	6	6

Table 15. Comparison of Samples Inducing Conception with Those Not Inducing Conception, Using Only Data on Initial Insemination

Criteria	Group I: Samples inducing conception N = 30	Group II: Samples not inducing conception N = 26	Diff. / S.E. of Diff.
Glycolysis per 0.1 ml. semen	27.7	20.2	1.63
Glycolysis per unit number of sperms	0.1263	0.1128	0.85
Sperms per 0.0001 cm. semen	222.8	181.3	1.82
Vesiculated sperm in 500	13.9	11.6	1.28
Sperm with normal form in 500	386	378	0.32

insemination, the variance of samples failing to induce conception is increased. Furthermore, it was evident that samples somewhat poorer than the average were good enough to have almost the same probability of inducing conception as the best samples used. Yet the criteria of whether fertilization did or did not occur in no way discriminated between two samples of different quality if both were good enough to cause fertilization. As a result of these two factors the ordinary method of comparing the difference between two means with its standard error is obviously not well adapted to the problem at hand. Another method was therefore resorted to. It consisted of ranking the samples relative to one of the criteria measured and then dividing them into consecutive groups of five samples each. The averages for groups of five of the criterion considered in ranking were then correlated with the number of the samples in the groups that induced conception. This was done for each criterion in turn. The correlations were calculated for the entire group of data and also for the smaller group consisting only of data on first inseminations. The object of this method of analysis was an attempt to correlate the level of samples with respect to a certain criterion with an experimental estimate of the probability of such samples inducing conception. The correlation coefficients are presented in table 16.

All correlations were positive but only the one with glycolysis per unit volume of semen in the data on first inseminations is

Table 16. Correlations between the Level of Various Characteristics of Semen and the Probability of Conception Resulting from Insemination with the Same Samples

Correlation between per cent conception and	All data 18 groups of 5 each	Data on first inseminations 11 groups of 5 each
Glycolysis per 0.1 ml. semen	.320	.681
Glycolysis per unit number of sperms	.386	.430
Sperms per 0.0001 cu. mm. semen	.266	.518
Vesiculated sperm in 500	.177	.110
Sperm with normal form in 500	.225	.424

significantly different from zero; none of the coefficients are different from each other by a significant amount.

The data from this experiment do not by themselves establish the relationship of any of the characters studied to semen quality; neither do they establish beyond any doubt which of these characters is most closely related to semen quality. However, considered together with other evidence they have some significance. Here, as in the case of Herman and Swanson *(25)*, the number of abnormal cells showed little relationship to whether or not individual samples induced conception. In contrast to the statement made by Walton and Edwards *(65)* regarding their data, there was an indication that the number of cells might be of some importance. As a result, glycolysis per unit volume of semen was correlated a little more closely to fertility than glycolysis per unit number of cells. While the correlations are not of sufficient size to be significant for the numbers involved, they support other evidence for the value of both glycolysis rates and vesicle number in measuring semen quality.

These results, together with those of Herman and Swanson, suggest that the relation between the quality of individual samples of semen and the probability of conception resulting from their use is obscured by the action of other factors affecting the occurrence of conception. Perhaps, therefore, more favorable results should not have been expected. There is some reason to believe, as indicated by table 14, that if samples of somewhat lower average quality had been used, the experiment might have been more decisive.

Correlations between Different Criteria of Semen Quality

Correlations between semen density, glycolysis rate per unit volume of semen, proportion of cells with vesicles, and proportion of cells with normal morphology calculated from two separate groups of data are presented in table 17. There is good agreement between the coefficients from the two sources. The important point is that glycolysis, vesicles, and abnormal cells are not highly correlated. Thus, if they are all important each must be measured in order to evaluate properly a sample of semen since any one of them gives very little information about either of the other two. An exception is the high negative correlation between the proportion of cells with vesicles and the proportion with abnormal heads when the average of the latter is high (Green and Comstock, *19*). This is doubtlessly due to the fact that vesicles are not found on cells with abnormal heads. The

Table 17. Correlations between Semen Density (z), Glycolysis Rate per Unit Volume of Semen (x), Proportion of Cells with Vesicles (u), and Proportion of Cells with Normal Morphology (t)

	Group I	Group II
Degrees of freedom	90	401
Simple correlations		
r_{xz}	.70	.60
r_{xu}	.35	.33
r_{xt}	.25	.22
r_{zu}	.34	.24
r_{zt}	.22	.31
r_{ut}	.24	.24
Partial correlations		
$r_{xu.z}$	.17	.24
$r_{xt.z}$	.14	.04

result is that high vesicular counts will seldom occur in samples with large numbers of head abnormalities, though the absence of the latter does not guarantee the presence of a large number of vesicles.

Sample Variation in Criteria of Semen Quality

The number of samples required to characterize the semen of a particular male depends on the variation among samples from the same animal. The within ram variances of table 18 were calculated from data on semen of 12 rams. With a few exceptions (occasionally a sample could not be obtained from one or another of the rams), 16 samples from each ram, four a month for four months, were involved.

The estimates of within ram variance were calculated on an intra-month basis in order to eliminate most of the seasonal effect on the characteristics. No attempt was made to remove the day-to-day effects of changes in weather since they would always be involved in an estimation of average semen quality. The mean

Table 18. Means, Ranges of Means for Separate Animals, and Sample Variation of Sperm Number per 0.0001 cu. mm. Semen (z), Glycolysis per 0.1 ml. Semen (x), Per Cent Normal Sperm (t)*, and Per Cent Sperm with Vesicles (u)*

	Characteristic			
	z	x	t	u
Variance	4460 ± 562	376† ± 48	.01566 ± .00201	.000411 ± .000052
Standard error	66.8	19.4	.125	.021
Mean	319	33.7	.688	.052
Mean monthly range of means for separate animals	236	41	.344	.054

* 500 cells observed.

† This is greater than a figure given by Comstock *(10)* for respiration rates which should have about the same variance as glycolysis rates. However, in that case the samples were collected on successive days instead of at weekly intervals through a month.

of all values used is given, since the variation is to some extent affected by the magnitudes measured. Variation in semen density and glycolysis rate decreases as the actual values become smaller. Variation in the proportion of normally formed sperm or of cells with vesicles decreases as the actual values approach either zero or one. The range of mean values of the separate animals was averaged for the four months and included in the table to give an idea of the differences that were encountered between animals. Except for density this range is between two and three times the standard error of individual samples. Thus a ram at either end of the range might produce a single sample falling almost anywhere within the range of the mean values. It should be stated that the average quality of the semen produced by the different rams varied from what would be considered very good to rather poor. Obviously more than one or two samples should be tested if accurate estimates of the average quality of semen of different animals are required. This is in accord with the recommendation of Herman and Swanson *(25)*.

The relative effects of the number of separate samples examined and number of cells observed per sample on the accuracy of estimating the proportion of cells with vesicles or of cells with normal morphology in the semen produced by different males are of practical importance. Error arises from (a) variation among different samples from the same animals and, (b) random distribution of cells of different types in a particular sample. More precisely:[6]

$$s^2 = \frac{s_1^2 + pq/N}{k}$$

where

s^2 is an estimate of the variance of determining the mean proportion of cells with normal morphology (or with vesicles) in the semen of a male,

s_1^2 is an estimate of the variance among the proportions occurring in different samples from the same male,

p is the mean proportion of cells with normal morphology (or with vesicles),

$q = 1-p$ is the mean proportion of cells with abnormal morphology (or without vesicles),

N is the number of cells observed per sample, and

k is the number of samples examined.

[6] Actually the expression given is only an approximation to the form,

$$s^2 = \frac{s_1^2}{k} + \frac{\sum_1^k \frac{(p_1 q_1)}{N}}{k^2}$$

where the values of p and q for individual samples are specified instead of the mean values for the male. The difference is not of sufficient magnitude to be important in connection with the matter at hand. In either case the assumption is made that the distribution of normal and abnormal cells on the microscope slide is random.

Obviously the examination of more samples decreases variance from both sources, while the observation of more cells per sample reduces variance from the latter only. Substitution of the experimentally estimated values of s^2, p, and q in the above equation gives

$$.01566 = s_1^2 + \frac{(.69 \times .31)}{500}$$ (Since the variance in table 18 is for single samples, $k = 1$.)

$s_1^2 = .01523$. (In the case of the proportion of normal cells)

The variance of determining the proportion in a particular sample, qp/N, is largest, for equal values of N, when $p = 0.5$. Therefore, by substituting 0.25 for pq in the formula, the error of determining the proportion in individual samples and, consequently, the value of increasing N is placed at a maximum. Using the above estimate of s_1^2, the standard errors of estimates of the mean proportion of normal cells can be approximated for different values of k and N. A series of such values was determined and appears as the first part of table 19. The second portion of the table is comprised of a similar series calculated in exactly the same way from data on bull semen presented by Herman and Swanson *(25)*. The variance between samples within males was based on only a slightly smaller number of degrees of freedom than in the case of the ram data presented by the authors and, as in that case, was calculated on an intra-month basis. The two series do not indicate exactly the same relative effect of changes

Table 19. Estimated Values of s for Different Values of k and N*

(N) Number of cells observed per sample	(k) Number of samples examined				
	1	2	3	4	5
Minnesota ram data					
25	.159	.112	.092	.079	.071
50	.142	.101	.082	.071	.064
100	.133	.094	.077	.067	.060
250	.127	.090	.074	.064	.057
500	.125	.089	.072	.063	.056
∞	.123	.087	.071	.062	.055
Data from Herman and Swanson† on bull semen					
25	.126	.089	.073	.063	.056
50	.105	.074	.060	.052	.047
100	.092	.065	.053	.046	.041
250	.083	.059	.048	.042	.037
500	.080	.057	.046	.040	.036
∞	.077	.054	.044	.039	.034

* s = estimated standard deviation of the mean proportion of cells with normal morphology among kN sperm cells.
N = number of cells observed per sample.
k = number of samples examined.
† Data from the first four months of 1940 in tables 6-13, inclusive, of their publication.

in the size of k and N, but the difference is not large enough to affect the general conclusions to be drawn. It is obvious from the table that the classification of more than about 100 cells per sample is an impractical refinement since it results in only a small reduction in the standard error of the estimate obtained. Important increases in the accuracy will result only from the examination of more samples per animal. Indeed, equal or greater accuracy can be attained from the observation of a smaller total number of cells by examining more samples but less cells per sample. In most work reported in the literature, 333 or 500 cells were observed per sample. This was entirely justifiable in the experimental stages of the work but should not now be recommended for practice in routine determinations.

The preceding discussion pertains only to the accuracy of determining the proportions of normal and abnormal (all types grouped) cells. The accuracy with which the proportion of a particular type of abnormality of low frequency is determined was not considered because, as mentioned earlier, there is no general agreement as to any one type of abnormality being more important than another in semen evaluation.

Using the same approach it is easily shown that 100 cells can be considered as a practical maximum to observe per sample when estimating the average proportion of sperm with vesicles in semen produced by a male.

Methods for Semen Appraisal

Optimum use of the various methods of semen appraisal depends on consideration of (a) the information given by each and (b) the relative importance of the various characteristics of semen in specific instances. It is important to remember that several characteristics may be necessary in sperm cells of high quality and that these characteristics may be relatively independent of each other, i.e., cells strong in one respect may be weak in another. If this is true *there probably is no single criterion aside from the breeding test that can be used as an accurate overall measure of sperm cell quality;* though a number may be useful measures of particular requisite characteristics. At the same time the relative importance of the several characteristics may shift, depending on the way in which semen is to be used, with the result that the information given by a particular method may assume a greater relative importance in one case than in another. In view of these considerations a brief summary of the type of information given by the various methods of evaluation appears

pertinent, though it will necessarily be incomplete in many respects because of lack of information.

Reaction rates (glycolysis and respiration) can be discussed together since they are closely correlated and because there is no evidence to indicate that one is superior to the other or that they yield separate types of information. They appear to be important specifically in the measurement of sperm cell motility and the ability of sperm cells to remain motile. On the other hand, the evidence indicates clearly that the ability to remain motile in storage can be determined more accurately by actual observation than it can be predicted from reaction rates. There is no evidence as to whether direct measurements of motility duration in storage or reaction rates are superior for the prediction of livability in the female reproductive tract; different factors may bear on the retention of motility in different environments. There is no particular reason to believe that reaction rates are closely related to important characteristics of sperm cells other than motility or the factors affecting motility.

The staining method of Lasley and co-workers is another technic relating to motility. Its use following temperature shock is reported of value for predicting motility duration. There is at present no evidence on the accuracy of this method compared to either visual estimates of motility or the measurement of reaction rates.

The remaining methods of evaluation have not been shown to be as specifically related to motility as the foregoing ones, nor has their value been claimed on that basis. Rather it must be concluded that they are related to one or more other necessary characteristics. The most widely used of these is the enumeration of sperm cells with abnormal morphology. Its usefulness appears to lie in the detection of males whose fertility is unsatisfactory in natural matings. While certain individuals are not correctly classified, these cases appear to be in a decided minority. Some reservation must be made with regard to its use for the detection of males whose semen is unsatisfactory for artificial insemination. No great number should be incorrectly indicated as unsatisfactory, but there might well be an increase in the number of unsatisfactory males that were not detected. This could result from the fact that the ability to retain motility in storage, a characteristic to which the proportion of abnormal cells is not closely related, would be considerably more important than in the former case. Strangely enough, neither the data presented herein nor those of Herman and Swanson *(25)* indicate this method to have

any particular worth for predicting the value for insemination of individual samples of semen.

Most of the remarks made regarding the enumeration of abnormal cells can be repeated with respect to the utilization of biometric constants calculated from measurements of sperm head lengths. It appears from the data of Savage, Williams, and Fowler *(59)*, however, that a somewhat greater proportion of unsatisfactory males would go undetected by this method. On the other hand, Williams and Savage *(72)* found one infertile bull not indicated as such by a study of sperm cell morphology but which would have been detected by knowledge of head length distribution. It might be concluded that this method could be used to supplement but not replace the enumeration of cells with abnormal morphology.

The vesicular structure has been shown to be a definitive feature of different samples of semen. It is lost as sperm cells age but the proportion possessing it is not highly correlated with viability of the sample of cells as a whole. On the other hand, there is evidence to indicate that this structure has, or is indicative of, a function not related to metabolic activity but perhaps equally important in making fertilization possible. There appears no reason to believe that presence or absence of the vesicle bears on the same characteristic as the proportion of abnormal cells. While vesicles were not observed on sperm cells with abnormal heads, the low correlation (table 17) between proportion of cells with vesicles and proportion of cells with abnormal forms indicates no great tendency for a lower proportion of vesicles on the cells with normal heads in a sample containing a large number of abnormal cells than in a sample with few abnormal cells.

It is apparent from the foregoing that unfavorable results from any one of the tests discussed may suffice to detect semen of low quality. On the other hand, favorable results from any one of the tests cannot be taken as a guarantee of high quality semen since the presence of only one desirable characteristic is in all probability not sufficient. It is clear that, at least in our present state of knowledge, several tests are required for a thorough measure of semen quality. However, there is no guarantee that the methods discussed cover all factors concerned in quality.

The relative importance of the various characteristics and consequently the emphasis to be placed on different measures depends to some extent on circumstances. For example, number of sperm cells produced may not be particularly important in a male used for natural mating in a small herd or flock of females if the

sperm cells he does produce are of good quality. By contrast, cell number would be much more important when it is desirable to use a male to the limit of his capacity as in the case of sires owned by an artificial insemination organization. In like manner, motility duration in storage would be much more important in the latter case; it may also be important if hand mating is practiced and matings are not made close to time of ovulation.

The choice of methods may also be altered by the purpose for which evaluations are required. Thus, in the evaluation of the semen of a prospective sire, motility duration in storage can be measured directly, whereas in the selection of individual samples to be used for a purpose requiring motility duration over a period of time, this would be impossible because the samples under consideration would be, for practical purposes, nonfunctional when the test was completed. In this case one of the methods for predicting motility duration in advance would be useful. To cite another example, if, as stated by Williams and Savage *(72)*, the proportion of cells with abnormal forms does not change during storage, that characteristic would not be useful for comparing methods of semen preservation.

A disadvantage of visual estimates of motility, the measurement of motility duration, and the enumeration of cells with abnormal morphology or vesicles results from the fact that they are partially subjective in nature, i.e., they involve a certain amount of judgment on the part of the technician. As a result, different workers would in all probability report somewhat different values for the same samples. Comstock *(8)* found this to be true in the case of motility estimates and also found that the workers differed in the relative evaluation of different samples. This would necessarily cause some disagreement between the results of different workers with respect to critical levels of these particular measures. What may be more important, an individual worker's judgment may shift from time to time. By comparison, measurements of reaction rates are relatively objective in nature. How important this factor is or whether it is of practical importance at all cannot be decided without information regarding the magnitude of its effect in practice.

The time and the amount and cost of special equipment required by the various methods and the ease with which they can be used in the field have not been considered, but are factors which will affect the selection of methods to be used for many routine purposes.

The measurement of reaction rates will usually be discrimi-

nated against the most because of equipment required, although respiration can be measured with apparatus (to be described in the next section) that is much less costly than the Warburg apparatus. A microscope and refrigerator are "musts" for work with semen and therefore cannot be regarded as special equipment.

The time required is much the greatest for the measurement of sperm head lengths and subsequent calculation of biometric constants, and is a serious disadvantage of that method.

The methods involving fixing and staining of the sperm cells are most easily adapted to use in the field. Suitable slides can be prepared on the farm or ranch with a minimum of special effort and examined at leisure in the laboratory at some later time. If sufficient samples are to be tested or the evaluation of one or two is of enough importance to make it practical, the portable apparatus to be described in the next section can be used to measure respiration rates in the field, provided electricity and a suitable working place are available.

It is obvious from the foregoing discussion that no set scheme for semen evaluation can be laid down for use in all circumstances. Rather, the technician must consider the importance of the decision to be based on the evaluation made, what information is most important to the decision, the amount of time and equipment it is practical to devote to the matter, etc., in choosing the method or combination of methods most appropriate to the particular case.

NEW EQUIPMENT FOR MEASUREMENT OF RESPIRATION OF SPERM CELLS

Extensive use of glycolysis or respiration rates for semen evaluation would require equipment which is both portable and moderate in cost. The Warburg apparatus does not meet these requirements but equipment that does has been devised in this laboratory. The principle involved is the same as for the Haldane gas analysis apparatus and the Barcroft apparatus modified according to Dixon *(13)*, namely, the direct measurement of volume change when pressure is constant. The reaction chambers are small flasks with calibrated capillary side arms open at the distal end. Measurements are made under water in a constant temperature water bath. Once temperature equilibrium has been reached, an amount of water equal in volume to the oxygen used by respiration is drawn into the side arm, where the quantity can be read directly.

Figures 3 and 4 show the essential parts of the apparatus. They are: (A) the units in which the reaction occurs and is measured, (B) a unit for making a blank determination, (C) a brass stand which supports the measuring units and holds them under water, (D) a clip which holds the blank unit, (E) a constant temperature water bath, and (F) a small electric motor with paddle for circulating the water in the bath.

Each measuring unit consists of a small flask (8-10 ml.) connected through a ground-glass joint and glass tube with a capillary side arm which projects at an angle of about 85 degrees with the perpendicular axis of the flask. A 0.2 ml. pipette (graduated to 0.001 ml.) attached with the zero mark at the distal end serves very well for the side arm since it furnishes a suitable volume calibrated to sufficiently small units. Results will be most satisfactory if the end of the pipette is removed about 2 mm. below the zero mark. If the end is then fire-polished, care must be exercised to avoid reducing the size of the opening since this results in irregular behavior. Each flask contains a small (0.3-0.4 ml.) inset cup. The volume of the flasks, including the one used for the blank determination, should vary as little as possible. Glass hooks should be fused onto the two parts of the ground-glass joints (the units in figure 3 lack them) so that rubber bands can be used to hold the joints firmly during use. The units in the figure have stopcocks at the proximal end of the side arms but these were found not to serve a practical purpose.

The unit for the blank determination had to be constructed so as to admit water part way into the side arm; this was necessary for the measurement of deviations resulting from contraction or expansion within the flask. As a result, this unit differs in two respects from the units described above. First, the tube leading from the flask is a T, one end of which connects to the side arm and the other through a stopcock to a 10-12 inch piece of rubber tubing. Second, there is a stopcock at the distal end of the side arm.

The construction of the brass stand (C) which holds the measuring units is evident from the figure. It is important that when the units are in position, the bases of the flasks shall be at least an inch above the base of the stand. Otherwise circulation between the flasks and the base of the stand is impeded and temperature equilibrium is reached too slowly when the stand and units are put into the water bath. The notched cross bar serves to keep the side arms in parallel position.

The water bath must be small enough to be portable but large

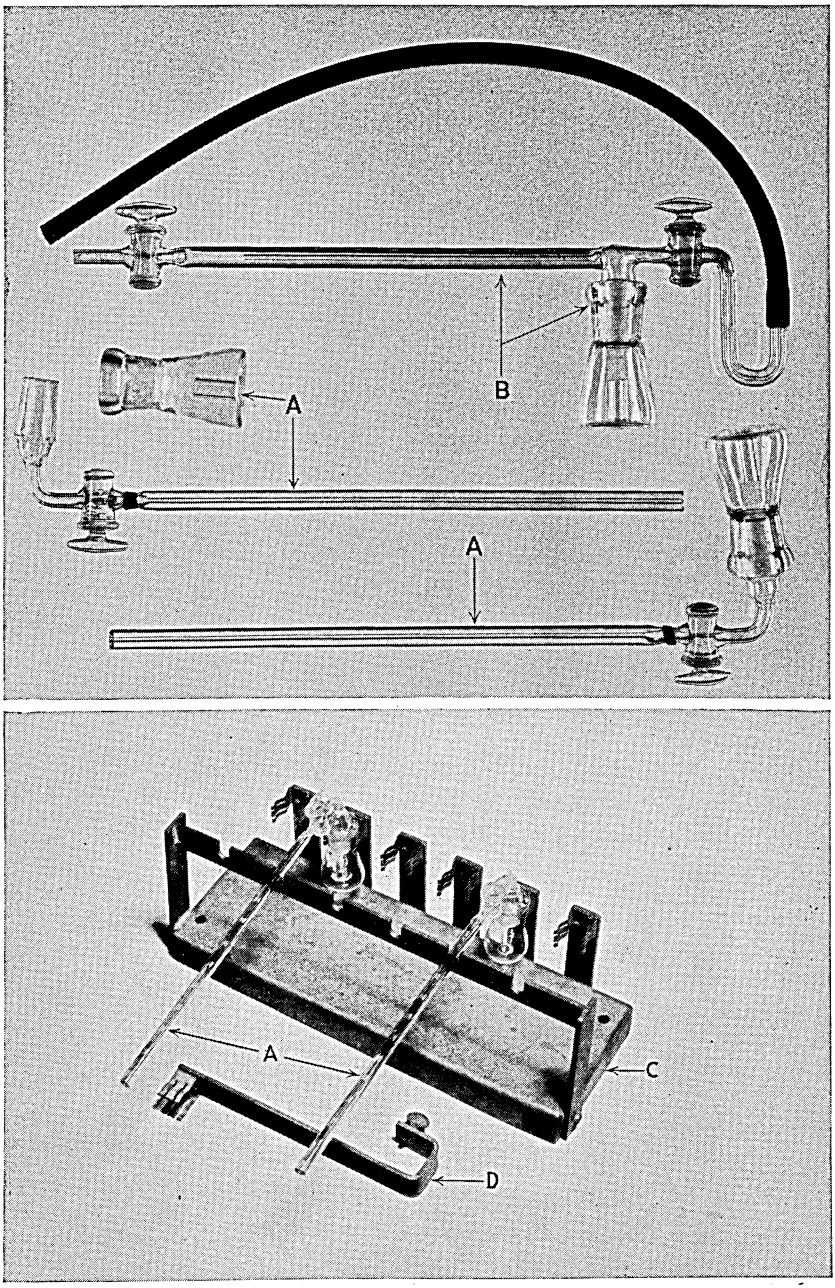

FIG. 3. Parts of the portable apparatus for measuring respiration of sperm cells (Above, 3A x ½, 3B x ¼)

(A) The measuring units, (B) Unit for making blank determination, (C) Supporting stand with two measuring units in position, (D) Clip for blank unit

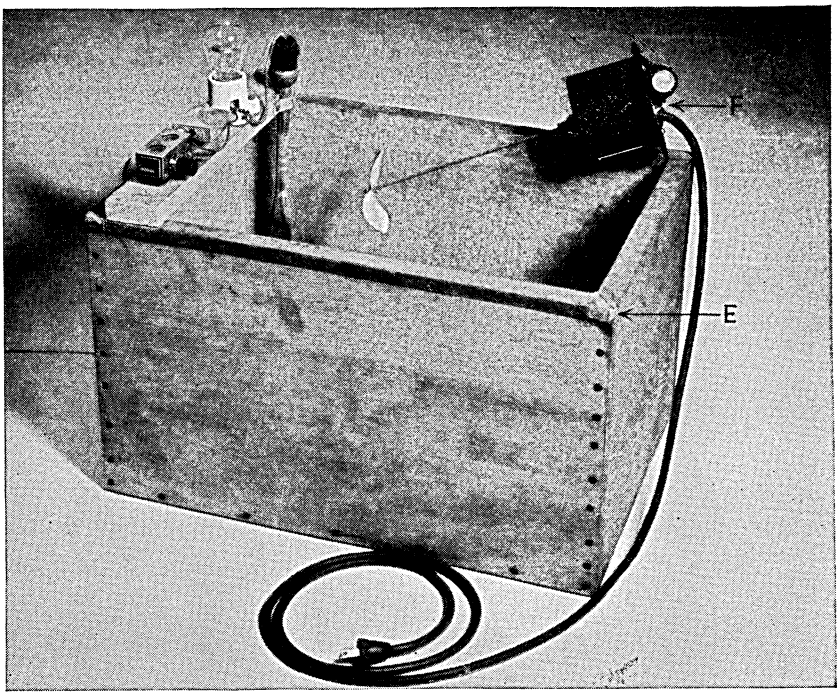

FIG. 4. Parts of the portable apparatus for measuring respiration of sperm cells *(E) Constant temperature water bath, (F) Electric stirrer*

enough to furnish sufficient working room. The one shown in figure 4 (12″ x 18″ and 12″ deep) was found to be satisfactory for use with six measuring units. It is constructed of pine board, lined with galvanized iron, and fitted with a stopcock for draining. Heat is supplied by a 250 watt knife edge heater (this could be replaced by a light bulb) wired in series with a relay and in parallel with a two watt pilot light. The relay is controlled by a thermostat (bimetallic helix type) on a current reduced by use of a fixed resistance.

Steps in the procedure of measuring respiration of sperm cells are as follows: The bath is filled with clean water of approximately the desired temperature and allowed sufficient time to reach the exact temperature for which the thermostat is set, e.g., 37° C. (98° F.). The unit for the blank determination is then prepared by drawing water up to the calibrated portion of the side arm by means of suction on the rubber tube, after which the stopcock on the side arm is turned off. With the other stopcock open, the unit is immersed and placed in the clip (D) which holds it under the surface of the water. The rubber tube must extend

outside the water bath allowing for expansion of the air within the flask as the temperature rises to that of the bath. After temperature equilibrium is reached (about 10 minutes), the stopcock leading to the rubber tube is closed and the other one opened. Water will then rise to about the center of the side arm, after which deviations in the water level will result mainly from changes in the temperature of the water bath. The loaded measuring units (semen and suspension fluid[7] in the main part of the flask, potassium hydroxide for absorption of carbon dioxide in the inset cups) with side arms attached are placed in the brass stand and set on the bottom of the water bath. After 10 minutes (water should by then have risen above the zero mark) readings of the water level in the side arms are made with the aid of a suitable light; the blank reading is made at the same time. At the end of a specified period of time (40 minutes has been used at Minnesota) readings are again made. Negative differences in blank readings are added to the difference between initial and final readings of each of the measuring units; positive ones are subtracted.

It is particularly important that the interior of the side arms be absolutely clean and dry before use. This can be accomplished by cleaning with acid cleaning solution, rinsing with distilled water, and drying with alcohol and ether.

A few trials may be necessary to find the position of the stand in the bath that promotes the most uniform results in all flasks. Avoid having the flasks close to the heating unit.

Accuracy

Error was determined from uniformity trials using semen from the same sample in all six reaction flasks. The same test was made using the Warburg apparatus. The average standard error of measurement for four such trials in which respiration ranged from 36 to 95 cu. mm. of oxygen in 40 minutes was 3.92 cu. mm. for the Warburg and 5.50 cu. mm. for the apparatus just described.

Respiration rates of 36 separate samples of semen were determined by each method in a second test of accuracy. The correlation between results from the two methods was 0.88. An indirect estimate of the error of the new apparatus, the square root of the difference between the squared standard error of predicting the value measured by the new apparatus from the value measured by the Warburg and the variance of the Warburg (3.92^2) obtained in the previous trial, was 8.4 cu. mm. In this second test

[7] The same as used for measurement of reaction rates in the Warburg apparatus.

the person measuring out the semen and operating the Warburg was doing this work for the first time. Since optimum accuracy is not attained without practice, the figure, 8.4 cu. mm., may represent something less than maximum accuracy.

Some doubt may be expressed over attempting to measure respiration without shaking to facilitate the diffusion of oxygen into the water. If diffusion rate of oxygen tended to inhibit respiration in the new apparatus it should have affected the higher rates the most and this would have resulted in a curvilinear relationship between values obtained by the two methods. There was definitely no indication of curvilinearity, from which it appears that rate of diffusion is not a factor within the range of respiratory rates measured (0-150 cu. mm. per hour). While the maximum rate that can be measured before it does become a factor has not been determined, values over 150 cu. mm. per hour per 0.1 ml. of semen are high even for ram semen. With semen of other animals, for which respiratory rates are rarely as high as for ram semen, it is advisable to use more than 0.1 ml. to insure that the oxygen used by most of the samples will be fairly large in comparison to the error of the method. The most suitable amount is such that the best samples will use around 100 cu. mm. of oxygen in 40 minutes.

In order to test whether the magnitude of errors made increased with that of the respiration rate measured, the correlation between the size of errors of estimating the Warburg values from the values measured by the new apparatus, disregarding signs, and the estimated Warburg values was calculated and found to be only 0.07. Apparently the size of errors made is independent of the rate of respiration measured.

Discussion

When using this apparatus for measuring respiration rates as a means of evaluating the semen of different males, the difference in accuracy between it and the Warburg will be negligible in comparison with the variation of samples from the same male. The sample variance of glycolysis rates, measured by the Warburg, was estimated as 376 (table 18); it would be about the same for respiration. Assuming 376 as the variance of respiration rates measured by the Warburg, four as the standard error of the Warburg, and 8.4 as the standard error of the new apparatus (with careful operation it may be lower), the variance of respiration rates of different samples from the same male measured by the new apparatus can be estimated as $376 - 4^2 + 8.4^2 = 431$. The

standard error of the mean of three samples tested with the Warburg would be $\sqrt{376/3} = 11.2$, and of the mean of four samples tested with the new apparatus would be $\sqrt{431/4} = 10.4$. It is apparent that the difference in accuracy of the two methods can be more than compensated by testing an additional sample of semen. Obviously, the difference in accuracy will be more important when the respiratory rate of specific samples is sought rather than when, as in the case just discussed, specific samples are being tested to obtain an estimate of the respiratory rate of all semen that a male may produce within a certain period of time.

In order to obtain absolute values (0° centigrade and 760 mm. pressure), corrections must be made for temperature, the pressure of the water above the side arms, and barometric pressure. The first two will be constant and, obviously, the uncorrected values will be as useful as corrected ones. Barometric pressure deviations of 10 mm. of mercury require a correction of only 1.3 per cent and thus the corrections ordinarily required are so small in magnitude as to be relatively unimportant.

To summarize, the equipment described for the measurement of respiration is portable, moderate in cost, and sufficiently accurate to be useful for many purposes.

Factors Affecting Semen Production

There is a great deal of variation in the characteristics of different samples of semen. Knowledge of the factors responsible should clarify the reasons for differences in fertility between sires and in different flocks and herds. A large portion of the total variation is between samples from different individuals. Data presented in this publication are typical in this respect of a considerable amount of data to be found in the literature. Such individual differences are doubtlessly partially genetic and partially due to differences in environment to which the animals are or have been subjected. The possibility of genetic differences receives support from the observation by McKenzie and Berliner *(33)* that under the same environment Shropshire and Hampshire rams exhibited different seasonal trends in the characteristics of semen and sperm cells produced.

Environmental factors which may affect semen production are temperature at which the testes are maintained, amount of exercise, nutrition, and frequency of ejaculation. Of these, scrotal temperature has probably received as much attention as any other. The work on that subject has been reviewed by Phillips

and McKenzie *(54)*. They also presented evidence on two rams which strongly suggested that the semen quality of rams in high condition and/or carrying long fleece may frequently suffer because of inability to maintain sufficiently low scrotal temperatures.

EXPERIMENT 1

Twelve ram lambs of the Shropshire breed were used to test the effect of differences in (a) the amount of fat carried by rams and (b) the frequency of ejaculation on the quality of semen produced. The experiment began in January when the rams were about 10 months old and was continued for a period of one year. Six of the rams were full fed on concentrate mixture[8] in addition to all the alfalfa hay of good quality that they cared for. The other six were limited to an amount of hay and concentrate mixture which sufficed to keep them healthy but in only what might be described as fair condition. There was a sharp contrast in the amount of fat carried by the two groups. Samples were collected once every two days from two rams of each group, once daily from two of each group, and twice daily from two of each group. Semen density, glycolysis rate, and proportion of cells with normal morphology were determined with a few exceptions on four samples from each ram in each month. The quality of semen produced by rams treated the same was sufficiently variable so that significant differences resulting from treatments could not be demonstrated either for the entire year or for any month taken separately. Part of the full-fed group produced a lower quality of semen than any of the other rams, particularly during the summer and early fall months. One was still producing poor semen in January when the experiment was terminated. On the other hand, two of the full-fed rams produced some of the best semen even during the summer when the effect of high condition on scrotal temperatures should have been most pronounced. McKenzie and Berliner *(33)* advanced the theory that the frequent observation of excessive condition and poor semen production in the same animal may result because both arise from a common cause, namely, a particular type of endocrine constitution, and not entirely because of a cause and effect relationship between high condition and low quality semen. If their theory is correct it may serve to explain the variable results obtained in this experiment where high or low condition was im-

[8] The same as specified in section on materials and methods.

posed on the animals regardless of their natural tendencies. On the other hand, in view of the known fact that high scrotal temperatures impair spermatogenesis, it appears likely that any condition which would reduce the ability of the body to give off excess heat would tend to reduce semen quality during hot weather. It must not be concluded that condition by itself has no effect on the semen produced. All that can be concluded is that if it does tend on the average to reduce semen quality, the effect is sufficiently variable so that it could not be clearly demonstrated in this experiment.

There were no significant differences in semen quality because of frequency of ejaculation. It is necessary to use extreme caution in drawing conclusions. A different type of experiment in which each individual was at different times subjected to the different amounts of use or the same type of experiment used to test the effect of other levels of use might have shown it to be an important factor. McKenzie, Miller, and Bauguess *(34)* showed that frequency of ejaculation affected seminal quality in boars, and Weatherby et al. *(67)* presented some evidence of the same nature in the case of bulls. It is generally accepted that when a certain frequency of ejaculation is exceeded, the quality of semen produced suffers. The maximum frequency not accompanied by a decrease in seminal quality doubtlessly varies between species, breeds, and individuals. The results of this experiment appear to warrant the conclusion that on the average young, healthy Shropshire rams are probably capable of ejaculation twice daily with no ill effect on the semen produced. This is in agreement with literature on this subject reviewed by McKenzie and Berliner *(33)*.

Because treatment differences were not significant, the data are not presented separately for the treatment groups. Instead, averages of the various characteristics are presented by months in table 20 to demonstrate seasonal trends.

The important feature is the decrease in all characteristics during the summer months. McKenzie and Berliner *(33)* report the same trend in sperm numbers and numbers of abnormal forms, though the low point occurs about a month earlier than in their data. Erb et al. *(15)* report the quality of bull semen is also lowest during the summer.

The monthly averages indicate lower quality of semen in September than later in the fall and that glycolysis rate is still low in October. Since there was considerable variation among the rams, the semen produced by some of them during the early

Table 20. Means by Months of Sperm Number per 0.0001 cu. mm., Glycolysis per 0.1 ml., and ~~Per Cent of~~ Sperm with Normal Morphology Per 500

Interval during which samples were tested	Sperm number per 0.0001 cu. mm.	Glycolysis per 0.1 ml.	Normal morphology
1/31-2/4	307	43	246
3/6 -3/9	360	38	221
4/4 -4/7	280	25	180
5/9 -6/1	367	45	193
6/12-7/5	210	21	132
7/13-8/7	91	9	124
8/14-8/28	102	16	237
9/5 -9/25	152	20	283
10/9 -10/30	183	13	367
11/6 -11/27	287	27	394
12/6 -12/28	333	27	371
1/8 -1/29	288	36	399

fall must have been very poor. As emphasized by Green *(18)*, this has an important bearing on the success of matings made early in the breeding season. Whenever it is desired to have the bulk of a lamb crop born early and within a relatively short period of time, it should prove especially worth-while to verify the fertility of the rams to be used.

EXPERIMENT 2

Subsequent to the completion of the above experiment a further test was made of the effect of high condition. As before, one group was full fed and the other fed to maintain only fair condition. The earlier experiment had shown that if a difference were to be demonstrated it would be desirable to use more rams. Therefore, a total of 17 rams was used. Only one factor, degree of condition, was tested. As a result there were twice as many degrees of freedom for experimental error as in the first experiment. Glycolysis rates and semen density were measured on four samples from each ram in May and three in September. This change was made because differences between the semen of different rams on the same treatment had been found large in comparison to differences between samples of semen from the same ram. Testing more than three or four samples from each ram at the beginning and end of the experiment would, therefore, have been an impractical expenditure of effort. September was chosen as the time of the final tests because in Minnesota practice it marks the beginning of the breeding season. The condensed data of this experiment are presented in table 21. At the time the final tests were made, the rams were placed in three classes, depending on their degree of condition (high, medium, and low),

Table 21. Glycolysis per 0.1 ml. and Sperm Number per 0.0001 cm.^{3} of Semen Produced by Rams Differing in Condition

Ram	Glycolysis per 0.1 ml.			Sperm number per 0.0001 cm.^{3}		
	May	September	Decrease from May to September	May	September	Decrease from May to September
			HIGH CONDITION GROUP			
33	102	91	11	470	434	36
54	26	30	−4	421	242	179
101	32	8	24	331	89	242
107	38	6	32	429	17	412
108	52	7	45	350	50	300
109	62	19	43	441	179	262
112	79	23	56	463	231	232
Mean			29.6			238
			MEDIUM CONDITION GROUP			
61	27	45	−18	231	359	−128
131	71	69	2	395	318	77
63	28	18	10	238	144	94
77	76	48	28	461	274	187
115	32	36	−4	203	259	−56
117	33	7	26	331	73	258
Mean			7.3			72
			LOW CONDITION GROUP			
43	70	58	12	475	362	113
67	41	36	5	381	222	159
110	33	39	−6	405	348	57
114	30	41	−11	341	304	37
Mean			0.0			92

by a competent sheep specialist who had no other contact with the experiment and who knew nothing about the quality of semen produced by the different rams. This is the basis of the classification in table 21. Only two of the full fed rams (61[9] and 131) failed to be put in the high condition group. None of those on limited feed were put in that group. Statistical analysis revealed the fact that the decreases in glycolysis rate and semen density from spring to fall were significantly greater for the high condition group than for the medium and low condition groups. There were individual differences in the degree of condition attained by rams within the group on limited feed which corresponded roughly with the decreases in glycolysis rates and semen density. This appears on the surface to verify the theory of McKenzie and Berliner *(33)*. However, the rams were hand fed individually. The intent was to give those on limited feed such amounts as would keep them in comparable condition. It is impossible to say whether the differences in condition which did

[9] One of the staff veterinarians expressed the opinion that this ram was infested with nodular worms which might explain his failure to gain high condition with unlimited feed.

arise were entirely the results of inherent differences in the rams or, in part, of failure of the caretaker to make perfect adjustments in amount of food offered the individual rams. Even assuming the former to have been the case, this experiment strongly supports the probability that, regardless of differences in endocrine constitution, high condition by itself has a deleterious effect on semen quality.

EXPERIMENT 3

A third experiment was designed to test whether length of fleece affects semen quality during the summer and early fall months. Glycolysis rates and semen density were measured on two samples of semen from each of eight rams on March 29 and March 31. The rams were then paired on the basis of average glycolysis rate; the two with the highest glycolysis rates were paired, the two with the next highest glycolysis rates were paired, and so on. One ram of each pair selected at random was then sheared. On May 2 and July 14 tests were again made. The four rams not sheared in the spring were sheared after the July tests. Final tests were made on September 29. Each test involved two separate samples from each ram. The condensed data reported in terms of differences between the members of the pairs are presented in table 22.

Table 23 contains analyses of variance of the differences recorded in table 22 and indicates that the size of the differences in both glycolysis rates and density varied significantly between

Table 22. Differences* in Quality of Semen between Rams Shorn in April and Rams Shorn in July

Pair	March 29	May 2	July 14	September 29
	GLYCOLYSIS PER 0.1 ML. OF SEMEN			
1	—33.9	—44.3	11.4	—26.5
2	1.1	14.9	19.4	8.3†
3	—0.7	—0.3	30.3	15.2
4	—1.8	24.3	31.1	6.2
Average	—8.8	—1.4	23.1	0.8
	SPERM NUMBER PER 0.0001 CMM OF SEMEN			
1	—118	—26	—87	—126
2	—177	—37	92	—49†
3	—53	38	237	61
4	—90	—24	276	92
Average	—110	—12	130	—6

* Positive differences indicate an advantage for rams shorn in April.

† One member of pair 2 was not tested in September. These values obtained by use of Yates' *(75)* formula for estimating the yield of a missing plot in replicated field experiments.

Table 23. Analyses of Variance in Testing Effect of Fleece Length on Quality of Semen

Source of variation	Degrees of freedom	Sum of squares	Mean square	F
		GLYCOLYSIS RATES		
Date of test	3	2,259.57	753.19	5.69*
Pairs	3	3,855.71		
Date x pairs	8†	1,059.18	132.40	
Total	14	7,174.46		
		SPERM NUMBERS		
Date of test	3	115,758	38,586	6.43*
Pairs	3	75,319		
Date x pairs	8†	47,985	5,998	
Total	14	239,062		

* Indicates $P < .05$ that F differs from 1.0 because of errors of random sampling.
† Degrees of freedom for "Date x pairs" are 8 instead of 9 since the one calculated value in table 22 is not an independent variate.

dates of test. Examination of table 22 shows that in July the quality of semen produced by the rams sheared in the spring was better in relation to those not sheared than in March and May but that an advantage had already been gained in May. However, the rams sheared in July had improved relative to their pair mates when the final tests were made in September. It is obvious that length of fleece is a factor affecting semen quality.

The results of the last two experiments indicate that both high condition and long fleece have a deleterious effect on semen quality, as measured by glycolysis rate and semen density, during the summer and early fall when quality is normally low anyway as a result of seasonal factors. Both probably operate to impede loss of body heat and thus bring about high scrotal temperatures known to be harmful to spermatogenesis. While it is probably true that certain breeds and individuals are, because of their physiological constitution, less likely to produce poor semen under these handicaps, it must be concluded that if high fertility is desired during the breeding season, excessive condition or length of fleece during the summer should be avoided. While experimental work has not been done on the effect of high condition in other species, the authors do not hesitate to recommend that breeding males of all species of farm animals be kept in no more than medium condition if maximum fertility is desired.

Summary and Conclusions

Estimation of semen density from light transmission measurements made with a photoelectric colorimeter is faster and more accurate than with the hemocytometer.

Estimation of semen density by comparison with opacity standards is sufficiently accurate to be valuable for field use.

Reaction rates (glycolysis and respiration) are useful measures of viability of ram sperm. If they are measured per unit volume of semen, a criterion of semen quality weighted by both metabolic activity and number of sperm is obtained.

There is no evidence to indicate that glycolysis and respiration rates are not equally valuable in the evaluation of ram semen.

Reaction rates per unit volume predicted motility duration more accurately than rates per unit number of sperm. The explanation of this fact was found in the various correlations that existed between reaction rates, sperm numbers, and motility duration.

Motility duration of sperm was not predicted from any combination of characteristics of the fresh samples as accurately as by direct measurement. However, useful predictions can be made in cases where for any reason direct measurement of motility duration is not practical.

Evidence is presented indicating that the vesicular structure may be an important attribute of high quality sperm even though its presence is not highly correlated with viability.

The vesicular structure was not found on sperm with abnormally formed heads. It appeared on sperm with tail abnormalities and sperm entirely normal in form with equal frequency.

The correlations among various characteristics of semen were so low that if each characteristic is important each must be determined for accurate semen evaluation.

The variation from one sample of semen to another from the same male within a period of a month is of such size that several samples must be tested to estimate accurately the fertility of a particular animal.

Comparison of sample variation in the proportions of sperm with normal forms and of sperm with vesicles with the error in determining these proportions for a single sample makes it clear that the observation of more than 100 cells per sample is impractical when either of these characteristics is being used to test male fertility.

Data on semen which did and which did not induce conception when used to inseminate ewes failed to furnish strong evidence on the actual or relative value of the characteristics concerned for the evaluation of semen quality. Glycolysis rate showed the highest correlation with the occurrence of conception; the proportion of normal cells the lowest.

Choice of the most useful method for semen evaluation depends on the purpose for which the evaluation is to be made. If the decision to be based on the evaluation is important, all possible information should be obtained.

New equipment is described for measuring the respiration rate of sperm cells. It is portable, moderate in cost, and its accuracy is reasonably close to that of the Warburg apparatus.

Data are presented on the seasonal trends of various characteristics of the semen of Shropshire rams. Quality is reduced on the average in the summer and early fall. Therefore, it should be particularly worth-while to test rams before using them in breeding operations at that time of the year.

Semen quality during at least the early part of the fall breeding season is reduced in rams carrying excessive condition or long fleece through the summer months.

Literature Cited

1. ASDELL, S. A., and SALISBURY, G. W. The viability of spermatozoa in the abdominal epididymus and the failure of motile sperms to fertilize ova. Amer. Jour. Physiol. **132**:791-795. 1941.

2. BATEN, WILLIAM DOWELL. How to determine which of two variables is better for predicting a third variable. Amer. Soc. Agron. Jour. **33**:695-699. 1941.

3. BERKSON, JOSEPH, MAGATH, THOMAS B., and HURN, MARGARET. The error of estimate of the blood cell count as made with the hemocytometer. Amer. Jour. Physiol. **128**:309-323. 1939.

4. BRADY, D. E., and GILDOW, E. M. Characteristics of ram semen as influenced by the method of collection. Amer. Soc. Anim. Prod. Proc. **1939**:250-254. 1939.

5. BREWSTER, J. E., MAY, RALPH, and COLE, C. L. The time of ovulation and rate of spermatozoa travel in cattle. Amer. Soc. Anim. Prod. Proc. **1940**:304-310. 1940.

6. BURBANK, R. C. The quantitative standardization of sperm suspensions by means of opacity. Quart. Jour. Expt. Physiol. **25**:393-397. 1935.

7. CHANG, MIN-CHUEH, and WALTON, ARTHUR. The effects of low temperature and acclimatization on the respiratory activity and survival of ram spermatozoa. Roy. Soc. London, Proc., Ser. B, **129**:517-527. 1940.

8. COMSTOCK, R. E. A study of the mammalian sperm cell. Ph.D. Thesis. University of Minnesota. Filed in library. Univ. of Minn. 1938.

9. ———— A study of the mammalian sperm cell. I. Variations in the glycolytic power of spermatozoa and their relation to motility and its duration. Jour. Expt. Zool. **81**:147-164. 1939.

10. ———— New portable equipment for measuring respiration of sperm cells. Amer. Soc. Anim. Prod. Proc. **1940**:216-220. 1940.

11. ———— and BRADY, D. E. A study of characteristics of normal and abnormal semen of the sheep. Amer. Soc. Anim. Prod. Proc. **1937**:233-238. 1937.

12. ———— and GREEN, W. W. Methods for semen evaluation. I. Density, respiration, glycolysis of semen. Amer. Soc. Anim. Prod. Proc. **1939**:213-216. 1939.

13. DIXON, MALCOLM. Manometric methods as applied to the measurement of cell respiration and other processes. University Press. Cambridge University. Cambridge, England. 1934.

14. DONHAM, C. R., SIMS, B. T., and SHAW, J. N. Fertility studies in the bull. II. The relation of the microscopic findings in semen to its fertility. Amer. Vet. Med. Assoc. Jour. **78**:665-680. 1931.

15. ERB, R. E., ANDREWS, F. N., and HILTON, J. H. Seasonal variation in semen quality of the dairy bull. Jour. Dairy Sci. **25**:815-826. 1942.

16. FISHER, R. A. Statistical methods for research workers. Ed. 6, 339 pp. Oliver and Boyd, Edinburgh (etc.). 1936.

17. GREEN, W. W. The chemistry and cytology of the sperm membrane of sheep. Anat. Rec. **76**:455-473. 1940.

18. ———— Seasonal trends of sperm cell types in sheep. Amer. Soc. Anim. Prod. Proc. **1940**:207-210. 1940.

19. ———— and COMSTOCK, R. E. Methods for semen evaluation. II. Sperm cytology in relation to viability. Amer. Soc. Anim. Prod. Proc. **1939**:217-219. 1939.

20. ———— and WINTERS, L. M. Studies on the physiology of reproduction in the sheep. III. The time of ovulation and rate of sperm travel. Anat. Rec. **61**:457-469. 1935.

21. ———— WINTERS, L. M., and COMSTOCK, R. E. Artificial insemination of farm animals. Minn. Agr. Expt. Sta. Bul. 336 revised, 1942.

22. GUNN, R. M. C. Fertility in sheep. Artificial production of seminal ejaculation and the characters of the spermatozoa contained therein. Austral. Council Sci. & Indus. Res., Bul. 94. 1936.

23. HAMMOND, JOHN. The effect of temperature on the survival in vitro of rabbit spermatozoa obtained from the vagina. Jour. Expt. Biol. **7**:175-191. 1930.

24. HENLE, GERTRUDE, and ZITTLE, CHARLES, A. Studies of the metabolism of bovine epididymal spermatozoa. Amer. Jour. Physiol. **136**: 70-78. 1942.

25. HERMAN, H. A., and SWANSON, ERIC W. Variations in dairy bull semen with respect to its use in artificial insemination. Mo. Agr. Expt. Sta. Res. Bul. 326. 1941.

26. IVANOV, ELIE. De la fécondation artificielle chez les mammiferes. Arch. des Sci. Biol. [Leningrad] **12**:377-511. 1907.

27. JONES, D. G., and LAMOREUX, W. F. Semen production of White Leghorn males from strains selected for high and low fecundity. Poultry Sci. **21**:173-184. 1942.

28. KELLEY, R. B. Studies in fertility of sheep. Austral. Council Sci. & Indus. Res. Bul. 112. 1937.

29. LAGERLOF, NILS. Ein Beitrag zur Sterilitat in der Rinderzucht. Ztschr. f. Zücht. Reihe B, **32**:47-93. 1935.

30. LARDY, HENRY A., and PHILLIPS, PAUL H. The effect of certain inhibitors and activators on sperm metabolism. Jour. Biol. Chem. **138**:195-202. 1941.

31. LASLEY, J. F., EASLEY, G. T., and BOGART, RALPH. Some factors influencing the resistance of bull sperm to unfavorable environmental conditions. (abstract). Jour. Anim. Sci. **1**:79. 1942.

32. ———— EASLEY, G. T., and MCKENZIE, F. F. A staining method for the differentiation of live and dead spermatozoa. I. Applicability to the staining of ram spermatozoa. Anat. Rec. **82**:167-174. 1942.

33. MCKENZIE, FRED F., and BERLINER, VICTOR. The reproductive capacity of rams. Mo. Agr. Expt. Sta. Res. Bul. 265. 1937.

34. ———— MILLER, J. C., and BAUGUESS, L. C. The reproductive organs and semen of the boar. Mo. Agr. Expt. Sta. Res. Bul. 279. 1938.

35. ———— and PHILLIPS, R. W. Determining fertility in the boar. Mo. Agr. Expt. Sta. Bul. 328. 1933.

36. ———— and PHILLIPS, R. W. Measuring fertility in the ram. A preliminary report. Amer. Vet. Med. Assoc. Jour. **84**:189-202. 1934.

37. MACLEOD, JOHN. The metabolism of human spermatozoa. Amer. Jour. Physiol. **132**:193-201. 1941.

38. ———— The effect of glycolysis inhibitors and of certain substrates on the metabolism and motility of human spermatozoa. Endocrinology **29**:583-591. 1941.

39. ———— An analysis in human semen of a staining method for differentiating live and dead spermatozoa. Anat. Rec. **83**:573-578. 1942.

40. MASON, L. W. Sterility with special reference to the spermatozoon. Amer. Jour. Obst. and Gynec. **17**:376-385. 1929.

41. MILLER, E. G., and KURZROK, RAPHAEL. Biochemical studies of human semen. III. Factors affecting the migration of sperm through the cervix. Soc. Expt. Biol. and Med. Proc. **28**:857-859. 1931.

42. MOENCH, G. L. A consideration of some of the aspects of sterility. Amer. Jour. Obst. and Gynec. **13**:334-345. 1927.

43. ———— The investigation of the relation of sperm morphology to fertility by means of microdissection. Amer. Jour. Obst. and Gynec. **18**:53-56. 1929.

44. ———— The technic of the detailed study of seminal cytology. Amer. Jour. Obst. and Gynec. **19**:530-538. 1930.

45. ———— Do sperm morphology and biometrics really offer a reliable index of fertility? Amer. Jour. Obst. and Gynec. **25**:410-413. 1933.

46. ———— A consideration of some of the aspects of sterility. Amer. Jour. Obst. and Gynec. **32**:406-415. 1936.

47. ———— and HOLT, HELEN. Sperm morphology in relation to fertility. Amer. Jour. Obst. and Gynec. **22**:199-210. 1931.

48. ———— and HOLT, HELEN. Biometrical studies of head lengths of human spermatozoa. Jour. Lab. and Clin. Med. **17**:297-316. 1932.

49. MOORE, BURTON H., and MAYER, DENNIS T. The concentration and metabolism of sugar in ram semen. Mo. Agr. Expt. Sta. Res. Bul. 338. 1941.

50. MUNRO, S. S. Motility and fertilizing capacities of fowl sperm in the excretory ducts. Soc. Expt. Biol. and Med. Proc. **33**:255-257. 1935.

51. NALBANDOV, ANDREW, and CASIDA, L. E. Ovulation and its relation to estrus in cows. Jour. Anim. Sci. **1**:189-198. 1942.

52. PHILLIPS, R. W. The physiology of spermatozoa. Amer. Soc. Anim. Prod. Proc. **1935**:222-235. 1935.

53. ———— and ANDREWS, L. N. The speed of travel of ram spermatozoa. Anat. Rec. **68**:127-132. 1937.

54. ———— and MCKENZIE, FRED F. The thermo-regulatory function and mechanism of the scrotum. Mo. Agr. Expt. Sta. Res. Bul. 217. 1934.

55. QUINLAN, J., MARE, G. S., and ROUX, L. L. The vitality of the spermatozoon in the genital tract of the Merino ewe, with special reference to its practical application in breeding. Union So. Africa Dept. Agr., Rpt. Dir. Vet. Serv. and Anim. Indus., **18**:831-870. 1932.

56. REDENZ, ERNST. Über Atmung und Glykolyse der saugetierspermatozoen. Second Internatl. Cong. for Sex. Res. Proc. 1930.

57. ———— Über den Spaltungstoffwechsel der Spermatozoen im Zusammenhaug mit der Beweglichkeit. Biochem. Ztschr. **257**:234-244. 1933.

58. SALISBURY, G. W., WILLETT, E. L., and SELIGMAN, JACK. The effect of the method of making semen smears upon the number of morphologically abnormal spermatozoa. Jour. Anim. Sci. **1**:199-205. 1942.

59. SAVAGE, A., WILLIAMS, W. W., and FOWLER, N. M. A statistical study of the head length variability of bovine spermatozoa and its application to the determination of fertility. Roy. Soc. Canada, Trans., Ser. 3, **21**:425-446. 1927.

60. ———— WILLIAMS, W. L., and FOWLER, N. M. A study of the head length variability of equine spermatozoa. Canad. Jour. Res. **3**: 327-335. 1930.

61. SCHOTT, RALPH GEORGE, and PHILLIPS, RALPH WESLEY. Rate of sperm travel and time of ovulation in sheep. Anat. Rec. **79**:531-540. 1941.

62. SEYMOUR, FRANCIS I., and BENMOSCHE, MOSES. Magnification of spermatozoa by means of the electron microscope. Amer. Med. Assoc. Jour. **116**:2489-2490. 1941.

63. SNEDECOR, GEORGE W. Statistical methods applied to experiments in agriculture and biology. Ed. 2, 388 pp. Collegiate press, Ames, Iowa. 1938.

64. TERRILL, CLAIR E. Comparison of ram semen collection obtained by three different methods for artificial insemination. Amer. Soc. Anim. Prod. Proc. **1940**:201-207. 1940.

65. WALTON, ARTHUR, and EDWARDS, JOSEPH. Criteria of fertility in the bull. I. The exhaustion test. Amer. Soc. Anim. Prod. Proc. **1938**: 254-259. 1938.

66. WARBURG, OTTO. Versuche an ubelebenden Carcinomgewebe. Biochem. Ztschr. **142**:317-350. 1923.

67. WEATHERBY, E. J., REECE, R. P., and BARTLETT, J. W. The ability of dairy bulls to withstand regular service for artificial insemination during one year. Amer. Soc. Anim. Prod. Proc. **1940**:224-229. 1940.

68. WHITE, W. E. The duration of fertility and the histological changes in the reproductive organs after ligation of the vasa efferentia in the rat. Roy. Soc. London, Proc., Ser. B, **113**:544-550. 1933.

69. WILLIAMS, W. L. The role of the male in the pathology of reproduction. Vet. Med. **27**:288-302. 1932.

70. WILLIAMS, W. W. Technique of collecting semen for laboratory examination with a review of several diseased bulls. Cornell Vet. **10**:87-94. 1920.

71. ———— and SAVAGE, A. Observations on the seminal micropathology of bulls. Cornell Vet. **15**:353-375. 1925.

72. ———— and SAVAGE, A. Methods of determining the reproductive health and fertility of bulls: A review with additional notes. Cornell Vet. **17**:374-385. 1927.

73. WINCHESTER, C. F., and MCKENZIE, FRED F. Influence of cell concentration on respiration rate of sperm. Soc. Expt. Biol. and Med. Proc. **48**:648-653. 1941.

74. WINTERS, L. M., COMSTOCK, R. E., COLE, C. L., GREEN, W. W., and BULIK, J. J. Artificial insemination of farm animals. Minn. Agr. Expt. Sta. Bul. 336. 1938.

75. YATES, F. The analysis of replicated experiments when the field results are incomplete. Empire Jour. Expt. Agr. **1**:129-142. 1933.